Appraisal of manufacturer's response to the ACD for:

Azacitidine for the treatment of myelodysplastic syndrome, chronic myelomonocytic leukaemia and acute myeloid leukaemia

Produced by West Midlands Health Technology Assessment Collaboration
Unit of Public Health, Epidemiology and Biostatistics
University of Birmingham
Edgbaston
Birmingham B15 2TT UK

Authors
Richard Edlin, Lecturer in Health Economics[2]
Jeff Round, Lecturer[2]
Sandy Tubeuf, Research Fellow[2]
Wendy Greenheld, Systematic Reviewer[1]
Anne Fry-Smith, Information Specialist[1]
Chris Hyde, Senior Lecturer[1]
Martin Connock, Systematic Reviewer[1]

[1]Unit of Public Health, Epidemiology and Biostatistics, University of Birmingham
[2]Academic Unit of Health Economics, Institute of Health Sciences, University of Leeds
Centre for Clinical Haematology, Queen Elizabeth Hospital, University Hospitals Birmingham NHS Foundation Trust

Correspondence to Martin Connock, Unit of Public Health, Epidemiology and Biostatistics, University of Birmingham

Date completed Oct 23 2009

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 08/84/01.

Declared competing interests of the authors: No competing interests – Hyde, Connock, Round, Tubeuf, Edlin, Fry-Smith, Greenheld, Craddock

Acknowledgements Professor C Craddock (Centre for Clinical Haematology, Queen Elizabeth Hospital, University Hospitals Birmingham NHS Foundation Trust) provided clinical advice and comments on the clinical aspects of the report.

CONTENTS

APPENDICES

1 Purpose of the appraisal

The purpose is to appraise the analyses and economic model submitted to NICE by the manufacturer subsequent to the ACD and contained in a "response document" (RACD) and additionally in a "clarification document" (C-RACD) that details further analyses undertaken at the request of NICE.

2 Summary of Content of RACD and C-RACD documents

Overall survival

The essential thrust of the RACD is that modelling of overall survival is best served by lognormal fits to observed data (study AZA-001) with extrapolation to 25 years adjusted for age-dependent mortality; when fed to the economic model as base case input the manufacturer proposes that this modelling generates the most plausible ICERs for the treatment comparisons. The table below summarises the manufacturer's base case ICERs submitted in the RACD.

Treatment option	**Cost per QALY gained**			
	Base-case results		**Base-case results with patient access scheme**	
	No Vial sharing	**Vial sharing**	**No Vial sharing**	**Vial sharing**
Pre-selected for best-supportive care				
Azacitidine BSC	£46,632	£43,744	£44,803	£42,641
Pre-selected for low-dose chemotherapy				
Azacitidine LDC	£39,714	£37,173	£38,105	£36,203
Pre-selected for standard-dose chemotherapy				
Azacitidine SDC	£36,591	£34,012	£34,959	£33,028

The use of the lognormal fit is a departure from the original submission (loglogistic fit). The RACD has employed data from an extension of the AZA-001 trial and this has generated different parametric fits (lognormal, loglogistic, exponential, Weibull, and Gompertz) to those presented in the original submission.

The manufacturer has justified the choice of lognormal fit from amongst the various options a] on the basis of AIC scores for "goodness" of fit for the five parametric models examined, and b] upon the shape of the observed survival curve of patients from a German MDS registry which has been termed "real life data".

In response to NICE's requests for clarification the manufacturer has presented additional analyses (in the C-RACD); these encompass ICERs derived using the following models for overall survival:

- Weibull fits with extrapolation adjusted for age-dependent mortality.
- Exponential fits with extrapolation adjusted for age-dependent mortality.
- Exponential fits for azacitidine patients and lognormal fits for control patients, each with extrapolation adjusted for age-dependent mortality.
- Overall survival for control patients based on the MDS German registry data with application of hazard ratios to obtain overall survival for patients treated with azacitidine.

The table below is taken from the C-RACD document and summarises the most relevant ICERs generated in the RACD and C-RACD submissions.

Curve fit selection		**Incremental cost per QALY gained (vs azacitidine) for each comparator [including vial-sharing]**		
Azacitidine	**CCR**	**BSC**	**LDC**	**SDC**
No patient access scheme				
Weibull	Weibull	£66,239 [61,350]	£51,471 [47,615]	£54,507 [49,059]
Exponential	Exponential	£70,674 [65,191]	£61,759 [56,482]	£65,019 [57,246]
Log-normal	Log-normal	£46,633 [43,744]	£39,714 [37,173]	£36,591 [34,012]
Exponential	Log-normal	£80,113 [73,486]	£63,983 [58,205]	£59,453 [53,204]
Düsseldorf MDS registry		£75,332 [69,315]	£61,561 [56,382]	£96,247 [79,730]
Patient access scheme (7% discount)				
Weibull	Weibull	£63,177 [58,547]	£49,030 [45,444]	£51,058 [45,991]
Exponential	Exponential	£67,203 [62,103]	£58,418 [53,510]	£60,098 [52,869]
Log-normal	Log-normal	£44,804 [42,118]	£38,105 [35,742]	£34,959 [32,560]
Exponential	Log-normal	£75,917 [69,755]	£60,325 [54,952]	£55,497 [49,685]
Düsseldorf MDS registry		£71,522 [65,926]	£58,282 [53,466]	£85,790 [70,430]

Key: BSC: best supportive care; CCR: conventional care regimen; LDC: low-dose chemotherapy; MDS: myelodysplastic syndrome; QALY: quality-adjusted life-year; SDC: standard-dose chemotherapy

Structural / functional changes to the economic model

Changes to the economic model are summarised in the RACD appendix. They include increased functionality for the following: control of costing sources for unit cost input; control of cost of weekend administration of azacitidine; alternative adverse event assumptions; removal of redundant sheets from the model; consistent labelling of overall survival curves.

Additional issues addressed or considered

The manufacturer's RACD has addressed further issues beyond overall survival and model structure. Those issues with a material influence on the manufacturer's input to the base case economic model included: survival in the AML state (original submission calculations now corrected); double counting of adverse events (original submission calculations now corrected). weekend administrative costs for azacitidine (original submission calculations modified); use of NHS 2009/10 tariff in the calculation of costs (original submission calculations modified) ; vial-sharing.

RACD issues considered but having no influence on the base case model input included: utilities for the model's health states; UK treatment patterns for MDS patients including the issue of exclusivity of treatment options.

The manufacturer has introduced two further commercial in confidence (CIC) economic considerations. These were the adoption of a patient access scheme that allows for 7% reduction in the acquisition cost of azacitidine

3 APPRAISAL

3.1 Overall survival

The first section of the RACD concerns the face validity of the modelled overall survival.

The manufacturer's time horizon of 25 years for economic analysis required modelling overall survival beyond the short term observed data of about 4 years. In the original submission there was a lack of face validity in the base case log-logistic model for overall survival (unrealistic numbers of MDS patients survived to become nonagenarians). The manufacturer has introduced two modifications to address this problem:

- The use of extended data from the trial AZA-001 with exploration of five parametric models (exponential, loglogistic, lognormal, Weibull and Gompertz) to fit the observed data.

- The adoption of a lognormal model (rather than log-logistic) with adjustment of the extrapolation of the parametric fit so as to allow for age-dependent mortality.

The selection of lognormal as the most suitable fit was firstly justified on the basis of AIC scores for "goodness of fit". However it should be noted that:

- There is no formal statistical test that allows comparison of different AIC scores.[1]

- "The choice of model may not be clear and supplementary information may be needed. For example comparison with other published results may be required to judge the relative plausibility of models rather than relying on AIC values alone."[1]

The new AIC scores for each parametric model were presented in the RACD appendix and are reproduced below.

Table A1: AIC values for curve fits to overall survival data including the AZA-001 extension data

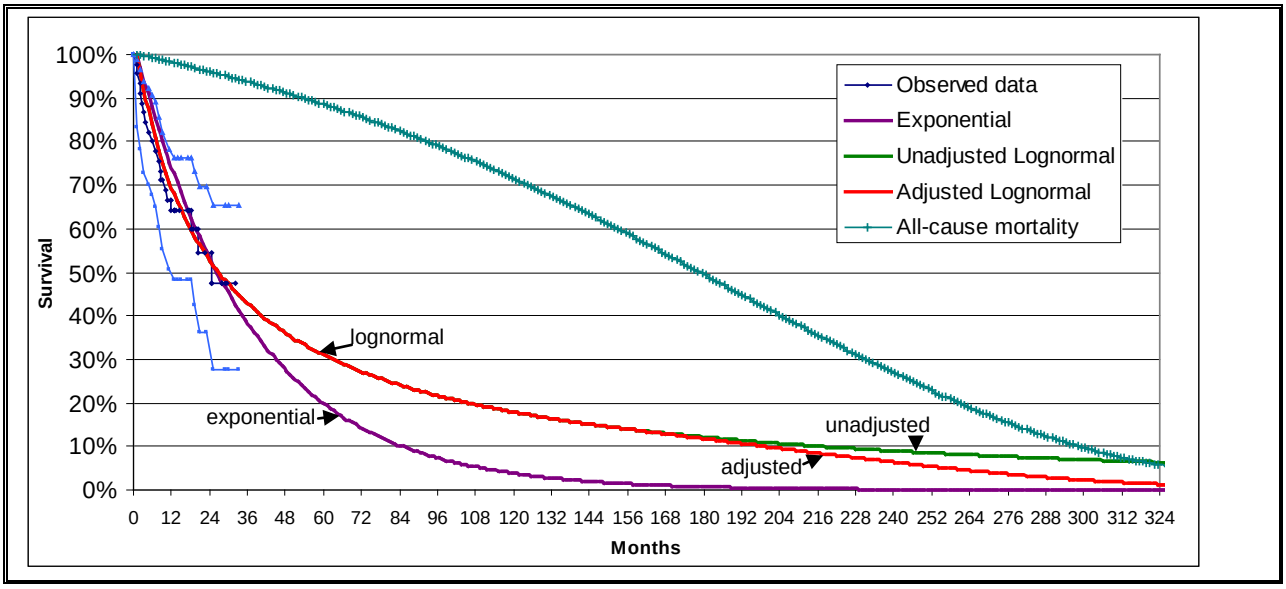
Fitted distribution	AIC for pre-selected subgroup					
	Azacitidine (BSC)	Azacitidine (LDC)	Azacitidine (SDC)	BSC	LDC	SDC
Exponential	301.2125	121.4813	48.85525	276.5794	130.7675	55.11062
Weibull	303.1845	122.7963	50.79613	277.3018	131.8855	51.44694
Gompertz	302.8256	122.0447	50.8036	278.464	132.1136	53.44262
Lognormal	303.6514	120.9462	51.00725	270.196	131.3413	50.24947
Log-logistic	302.7885	121.9108	50.88705	271.382	132.7571	51.19378

COMMENT

It should be noted that the exponential fit has the “best” score for three subgroups and the lognormal the “best” for three subgroups, furthermore (as acknowledged in the RACD) the AIC scores within each subgroup are very similar. This implies that, on the basis of AIC scores, there is little to distinguish between any of these fits and in particular between exponential and lognormal. As acknowledged in the RACD other considerations should be brought to bear, one of which is the biological plausibility of the extrapolated survival curves.

Because of the lack of face validity in the extrapolation of the original lognormal parametric fit the manufacturer has adjusted the extrapolations to include all-cause age-dependent mortality.

The effect of the adjustment on the lognormal extrapolation for the AZA-treated LDC-preselected subgroup was illustrated in RACD Fig 1A (and Fig A1c) shown below.



COMMENT

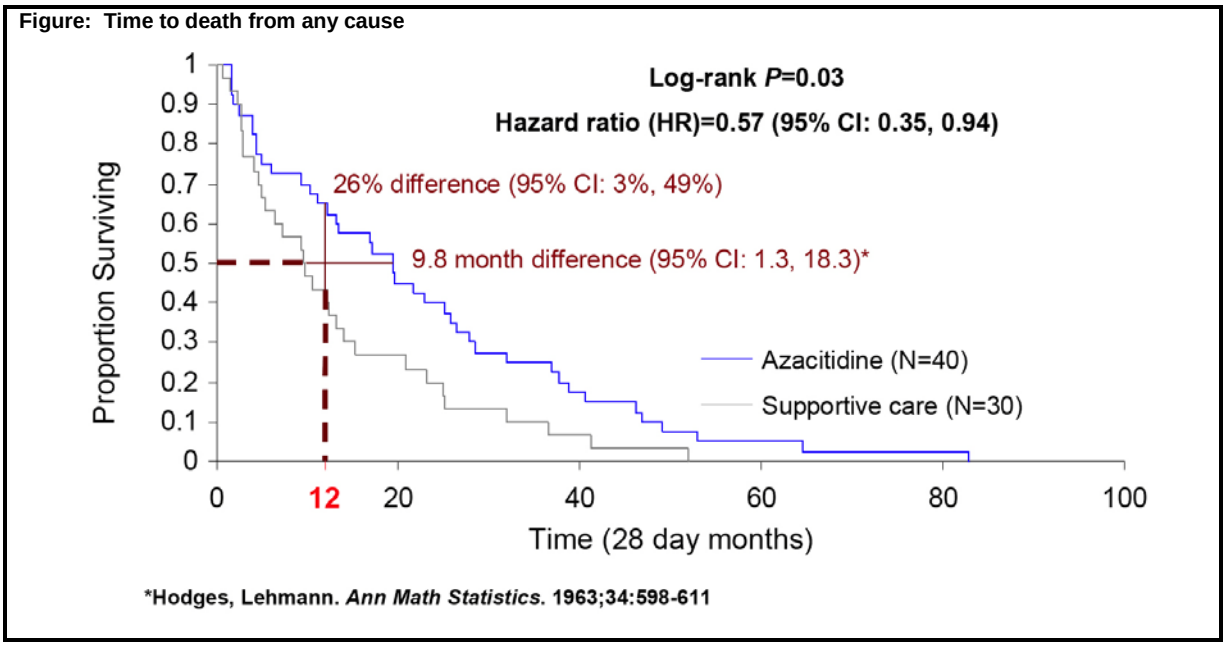
The observed data analysed only appears to extend to less than 36 months; this is surprising given that the AZA-001 extension included 45 x 5-week cycles (4.3 years); it is possible the AZA-LDC subgroup were late entries into the trial or possibly the fits have been plotted onto old observed data. The latter possibility is disturbing because it is then unclear if the adjusted and unadjusted fits illustrated actually correspond to old or extended AZA-001 data and whether they correspond to the AIC scores tabulated in the RACD. Furthermore

there is confusion due to a mismatch in the time axes between the graphs for overall survival in the model (5-week cycles) and those in the RACD (months), see Appendix 1 for details. It should be noted that the extension data only applies to AZA-treated groups (see Appendix 2).

In the unadjusted lognormal model about 6 to 7% of patients are alive after 25 years (mean age for this subgroup at start was 69) yielding patients aged about 94 years. With adjustment the percentage of nonagenarians is reduced to about 1%. For the other AZA subgroups (see RACD page 28 figs A1a and A1c) the adjusted lognormal extrapolation yields about 1% (BSC-preselected group) and 2% (SDC-preselected subgroup) nonagenarians. These lognormal survivals are not compatible with the results of the CALBG 9221 study; in particular:

- At 72 months (6 years) in the three AZA-001 trial AZA subgroups about 27% patients are alive; this contrasts with the study CALBG 9221 in which all high risk AZA-treated patients were dead by about 6.4 years (83 x 28 day-months) as illustrated in Fig A2.1 of manufacturer's 8 April 2008 response for clarification and reproduced below.

Figure: Time to death from any cause



*Hodges, Lehmann. *Ann Math Statistics*. 1963;34:598-611

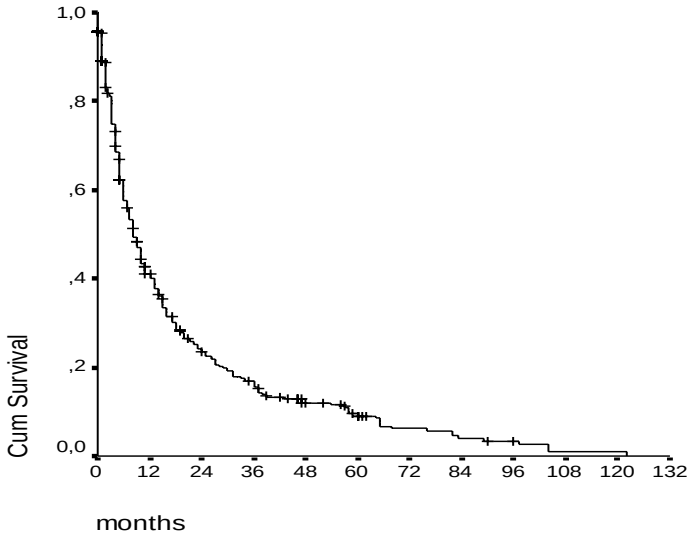
NICE requested clarification regarding Weibull, exponential and lognormal parametric models for overall survival. The C-RACD document provided appropriate graphs to illustrate these (see C-RACD figs 1.1 to 1.9). A feature of the lognormal extrapolations for the AZA-treated subgroups not shared by Weibull or exponential models is the predicted large

proportion of survivors beyond 12 years (10% to 20%) that is incompatible with results from study CALBG 9221.

To address the face validity of the flat tail of the survival curve seen with the lognormal models the manufacturer sought external data.

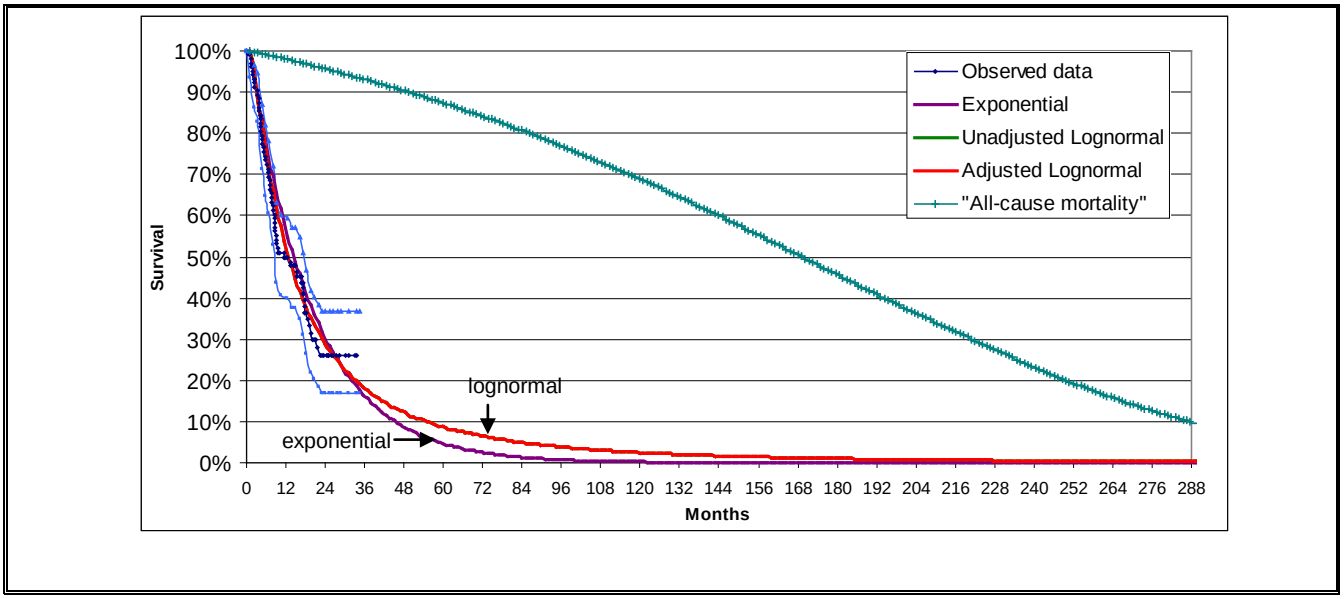
Celgene have also sought an external data source to further assess and present to the Committee the potential long-term survival for patients with high-risk MDS.

The external source used was a German registry describing survival of 655 high-risk MDS patients with mean age 70 years (range 18 – 96) treated only with BSC. The Kaplan-Meier curve for overall survival is shown below (RACD Fig A2).



COMMENT

The curve has a flat tail of long term survival extending from ~60 to 120 months, with no survivors beyond about 120 months (10 years). The RACD compares this registry curve with the lognormal (adjusted) and exponential fits to all AZA-001 trial BSC patients (RACD Fig A3 and also RACD FigA1d shown below).



COMMENT

- The adjusted AZA-001 lognormal fit (upper line) indicates ~ 3% survivors at 10 years with some surviving beyond this time (this is dissimilar to the registry curve). The exponential indicates few survivors beyond 108 months again dissimilar to the registry data. Relative to the registry data the lognormal fit appears overgenerous while the exponential is under-generous. Similarly the Weibull fit (C-RACD Fig1.1) is also under-generous.

- In the CALBG 9221 study all BSC patients were dead before 55 months.

The ERG also noted that the lognormal parameters in the model for both BSC and LDC subgroups are identical (table copied and pasted from the appropriate model sheets):

LDC subgroup	parameter	BSC subgroup
2.3831800	mu	2.3831800
1.1458937	sigma	1.1458937

Since the observed survival for the groups differs the ERG extracted BSC and LDC survival data from the model and used STATA software to obtain lognormal fit parameters. The output is summarised below:

GROUP	N	Failures	mu	sigma
LDC	49	31	**2.446731**	**1.188967**
BSC	105	66	**2.38318**	**1.145894**

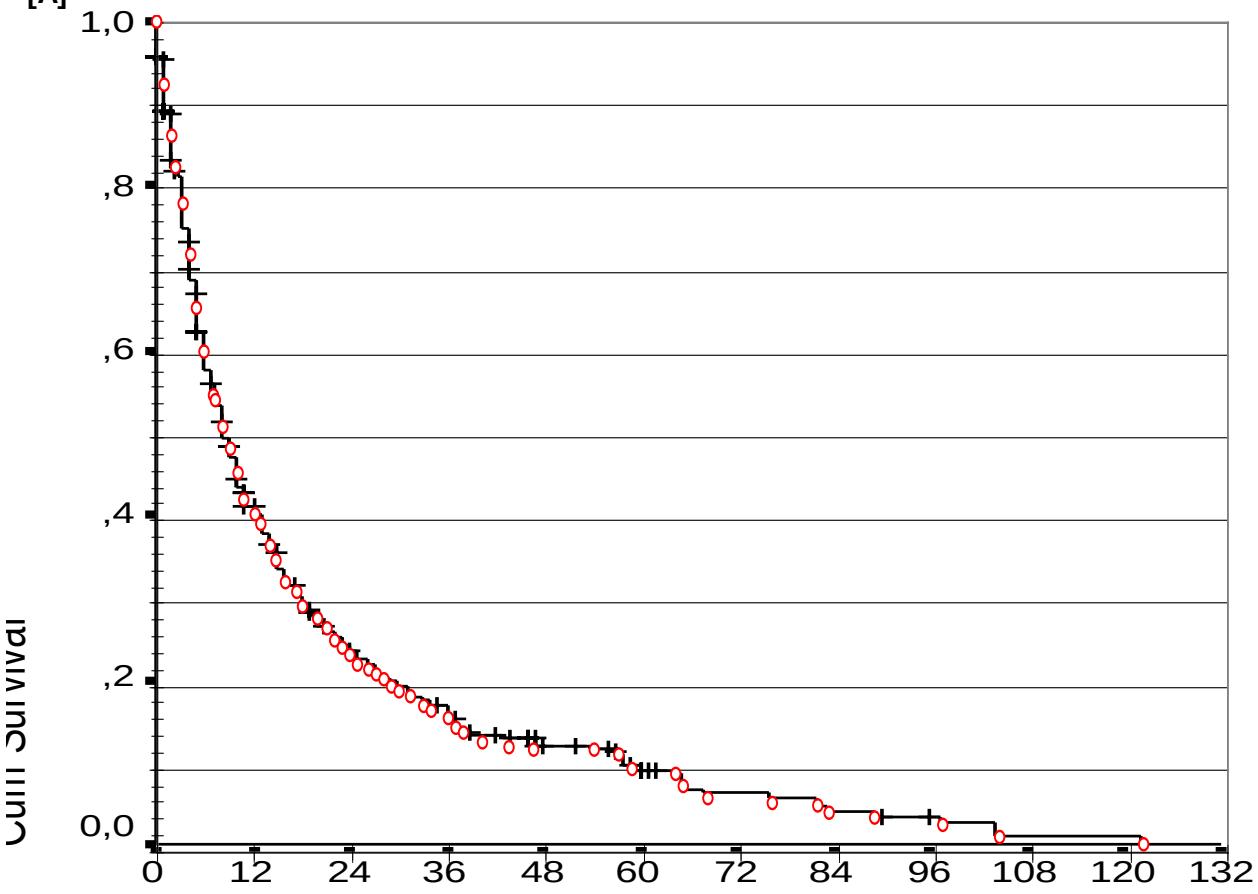
From this the ERG conclude that the lognormal parameters for the BSC subgroup have been entered into the model for both BSC and LDC subgroups.

On examination of parameters in the model for the other fits to observed survival for the BSC and LDC subgroups it appears that the same Gompertz parameter values have been entered for the BSC and LDC groups (but not for Weibull, exponential or loglogistic fits; see Appendix 3).

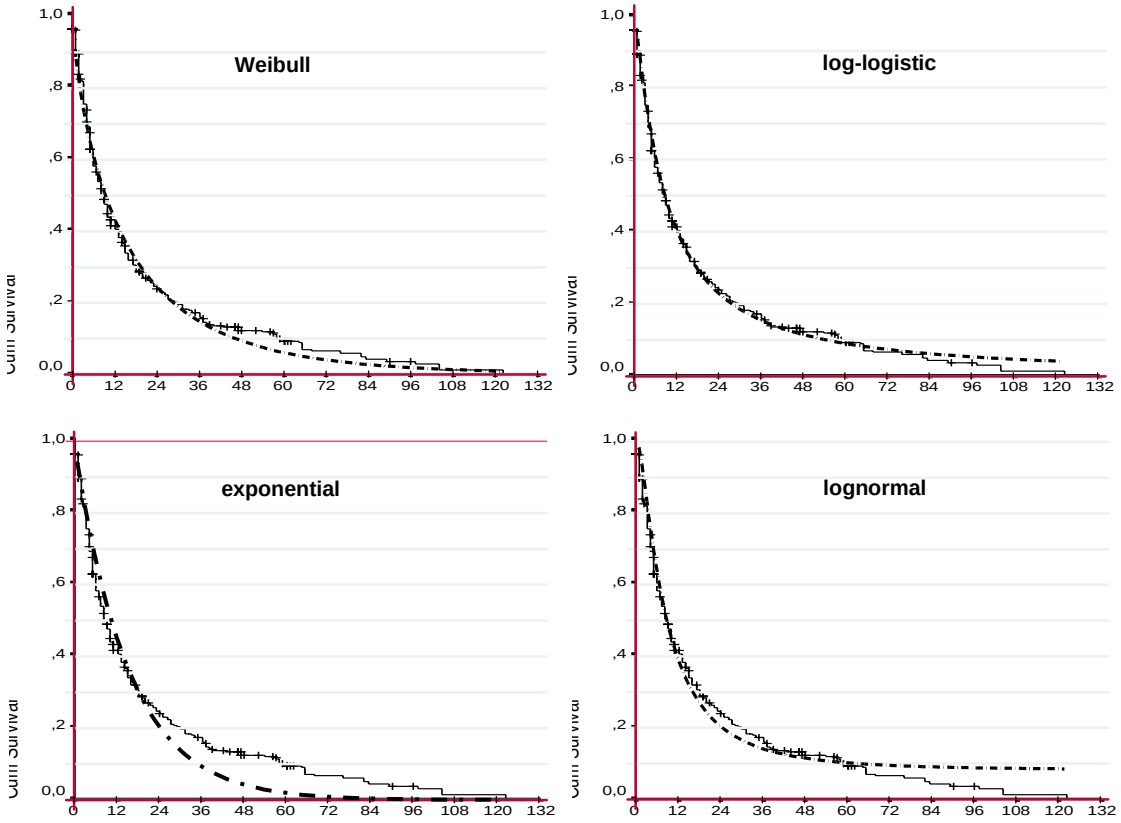
Parametric fits to registry survival data

One indicator of which parametric fit is most appropriate would be to compare the observed registry survival curve with its various parametric fits. Ideally this is done using individual patient level data. The IPD data could not be found in the model submitted by the manufacturer. The ERG therefore extracted data from the Kaplan-Meier registry graph and then generated parametric fits. The figures below show: [A] the correspondence between extracted data (dots) and registry plot; [B] parametric fits to the extracted data superimposed on the Kaplan-Meier plot for observed survival.

[A]



[B]



All four parametric models fit well to the early part of the registry Kaplan-Meier plot. With respect to the tail of the Kaplan-Meier, where the long-term survivors are represented, the parametric fits differ. The lognormal fit is flattened beyond about 72 months and implies many survivors beyond 120 months that is incompatible with the observed data. The log-logistic model also exaggerates the long term survivors relative to the "real-life data" although to a lesser extent than does the lognormal. The exponential fit appears to underestimate long term survivors while the Weibull arguably provides the best fit to the tail of the Kaplan-Meier indicating a small proportion of survivors to 120 months that is compatible with the registry data.

In summary:

1. The AIC scores provide meagre guidance regarding which parametric fit best describes observed overall survival in the AZA-001 study.

2. The AIC scores do not provide convincing evidence that lognormal is the most appropriate model for extrapolation of observed survival to 25 years.

3. The adjusted lognormal extrapolation for AZA-treated patients may still be considered moderately implausible in generating unrealistic proportions of nonagenarians.

4. The overall survival of AZA-treated patients that is predicted by the adjusted lognormal model is much greater than, and incompatible with, that observed in study CALBG 9221.

5. The adjusted lognormal fit for AZA-001 study BSC patients provides slightly better long-term survival than that observed for BSC patients in the German registry, and much greater than long term survival for BSC patients seen in the CALBG 9221 study. Weibull and exponential fits for AZA-001 BSC patients provide worse survival relative to registry BSC patients but are reasonably compatible with BSC patients in study CALBG 9221.

6. A lognormal fit to the German registry data for BSC patients generates a proportion of long term survivors that is incompatible with the observed data. Of Weibull, lognormal, log-logistic and exponential fits to the German registry data the Weibull model best describes the proportion of long term survivors while the lognormal appears the least appropriate.

Conclusion regarding modelling of overall survival

The choice of lognormal in preference to Weibull or exponential models to describe overall survival is not strongly supported by the available evidence. Weibull or exponential based models are at least as equally plausible as lognormal, and in the case of the Weibull are probably more plausible. NICE requested the manufacturer conduct economic analyses additional to the maunufacturer's lognormal base case and using several plausible models for overall survival. As discussed in a following section the lognormal model delivers ICERs between 23 and 33% lower than any of the other plausible models.

3.2 Functionality and validation of the submitted economic model

Model dated 07/09/09

A version of the excel-based model provided by the manufacturer incorporated a number of changes as requested by the ERG. A number of other requested changes were not made, and the manufacturer provides a series of reasons for why these were not included. These changes were outlined in the manufacturer's response to the ACD (RACD). Further details of changes and results of additional analyses were provided in the manufacturer's clarification to the response to the ACD (C-RACD) dated 07-09-2009.

On examination of the most recent excel model (dated 07-09-2009), a serious flaw was noted which called into question the reliability of any of the manufacturer's results that were based on this model. Below we examine where this error occurred and how it affects the results. We then, as an addendum to this report, include a brief overview of an earlier version of the model (dated 24-08-2009) and consider whether the results from this version can be considered sufficiently reliable for the committee to consider a decision based on them.

The flaw in the model was a simple typing error in a single cell. The error was potentially easily corrected. However, its impact on the results that were generated when running the model were significant.

The error and it's correction (provided by Celgene, received by the ERG on 22nd October, 2009).

The cause was located in the Input worksheet, cell CL56:
"=IF(VLOOKUP($CK56,range_StaffVidaza,4,1)=0,**"-"**,VLOOKUP($CK56,range_StaffVidaza,4,1))*IF(r_WeekEndCost="Yes",2/7*v_WEpharmacist+5/7,1)"

This has been amended to:
"=IF(VLOOKUP($CK56,range_StaffVidaza,4,1)=0,**0**,VLOOKUP($CK56,range_StaffVidaza,4,1))*IF(r_WeekEndCost="Yes",2/7*v_WEpharmacist+5/7,1)"

The correction to the model provided by Celgene was tested and a full set of results were generated. These have not been appraised in detail due to the late nature of the discovery of the error and the receipt of the correction from Celgene.

Provided below are screen shots of the model as received by the ERG and then again after we attempted to test the validity of the model and then learned that there was an error that

needed to be corrected. Although the details cannot be seen clearly, what is obvious is that in Figure 1 , there are a full set of results that can be viewed. After having checked the model to establish whether or not the changes claimed by the manufacturer had indeed been made (see the addendum to this report for more details) the ERG tested the internal validity of the model by running the model under the probabilistic setting. A screenshot of the results is presented in Figure 2 .

Figure 1 Results tab in model as received by the ERG.

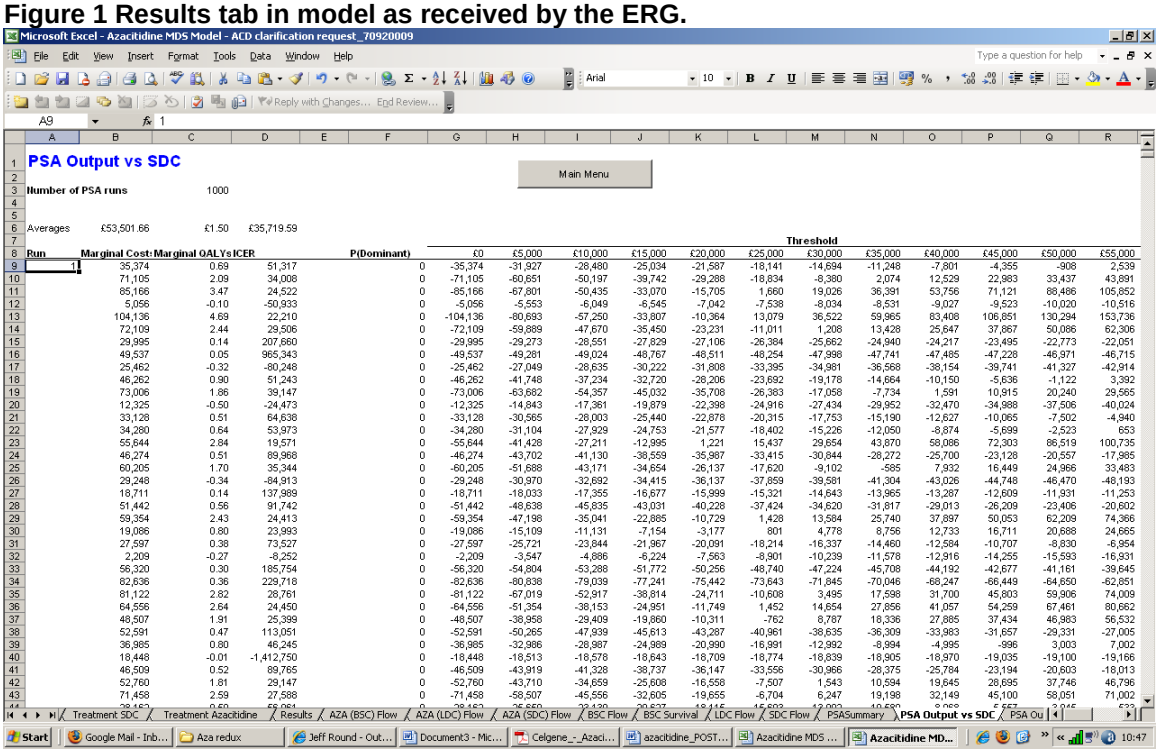
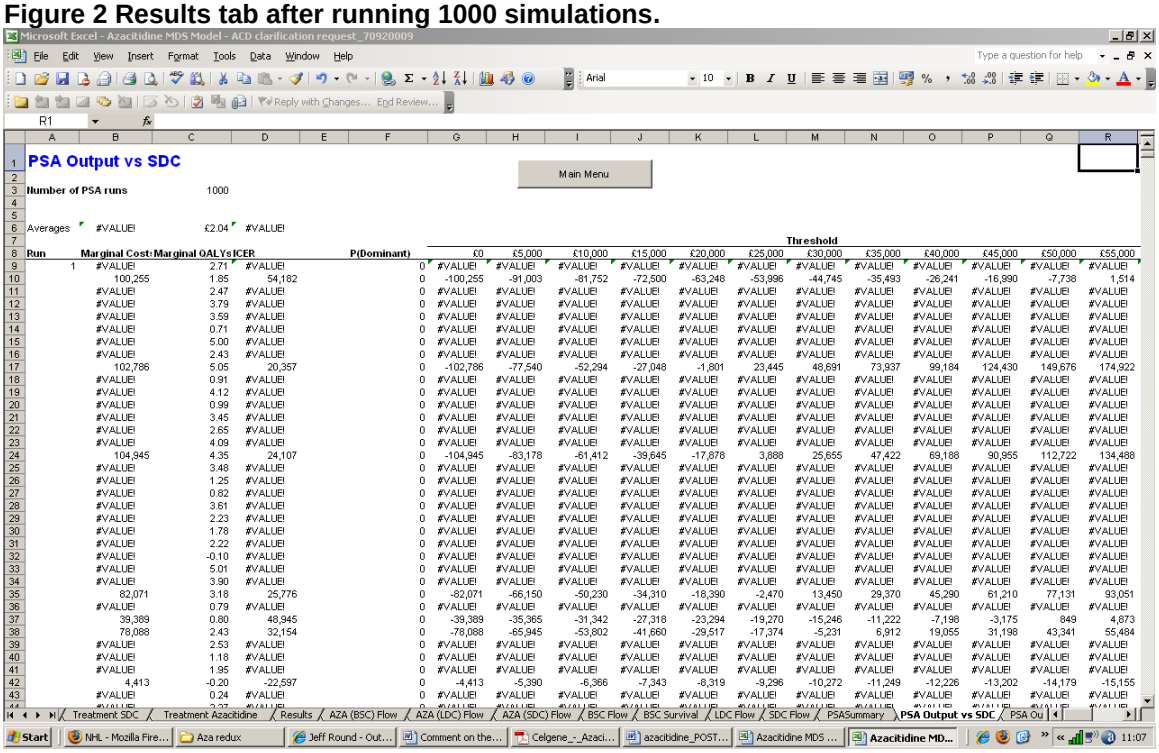


Figure 2 Results tab after running 1000 simulations.



PSA Output vs SDC

Main Menu

Number of PSA runs 1000

						Threshold											
Run	Marginal Cost	Marginal QALYs	ICER	P(Dominant)		£0	£5,000	£10,000	£15,000	£20,000	£25,000	£30,000	£35,000	£40,000	£45,000	£50,000	£55,000
1	#VALUE!	2.71	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	100,255	1.85	54,182	0	-100,255	-91,003	-81,752	-72,500	-63,248	-53,996	-44,745	-35,493	-26,241	-16,990	-7,738	1,514	
	#VALUE!	2.47	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	3.79	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	3.59	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	0.71	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	5.00	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	2.43	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	102,786	5.05	20,357	0	-102,786	-77,540	-52,294	-27,048	-1,801	23,445	48,691	73,937	99,184	124,430	149,676	174,922	
	#VALUE!	0.91	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	4.12	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	0.99	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	3.45	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	2.65	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	4.09	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	104,945	4.35	24,107	0	-104,945	-83,178	-61,412	-39,645	-17,878	3,888	25,655	47,422	69,188	90,955	112,722	134,488	
	#VALUE!	3.48	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	1.25	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	0.82	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	3.61	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	2.23	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	1.78	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	2.22	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	-0.10	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	5.01	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	3.90	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	82,071	3.18	25,776	0	-82,071	-66,150	-50,230	-34,310	-18,390	-2,470	13,450	29,370	45,290	61,210	77,131	93,051	
	#VALUE!	0.79	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	39,389	0.80	48,945	0	-39,389	-35,365	-31,342	-27,318	-23,294	-19,270	-15,246	-11,222	-7,198	-3,175	849	4,873	
	78,088	2.43	32,154	0	-78,088	-65,945	-53,802	-41,660	-29,517	-17,374	-5,231	6,912	19,055	31,198	43,341	55,484	
	#VALUE!	2.53	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	1.18	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	#VALUE!	1.95	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
	4,413	-0.20	-22,597	0	-4,413	-5,390	-6,366	-7,343	-8,319	-9,296	-10,272	-11,249	-12,226	-13,202	-14,179	-15,155	
	#VALUE!	0.24	#VALUE!	0	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!

Although the image quality is low, what can be clearly seen is that the majority of cells contain no values, as a result of the programming error. In this sample of 1000 simulations, the total number of valid simulations was 292. All other simulations returned an error message. The results shown above are for the comparison of azacitidine with standard dose chemotherapy (SDC). The same error applies throughout all comparisons made in the model.

Additionally, since the model as received by the ERG was ultimately shown to be non-functional, the ERG have questions about how it came to be that the model version 07-09-2009 included a full set of results when opened. It is clear that these could not have been generated by the model having been run, as the error would not permit it. It also calls into question where the results presented in the RACD and the C-RACD were obtained.

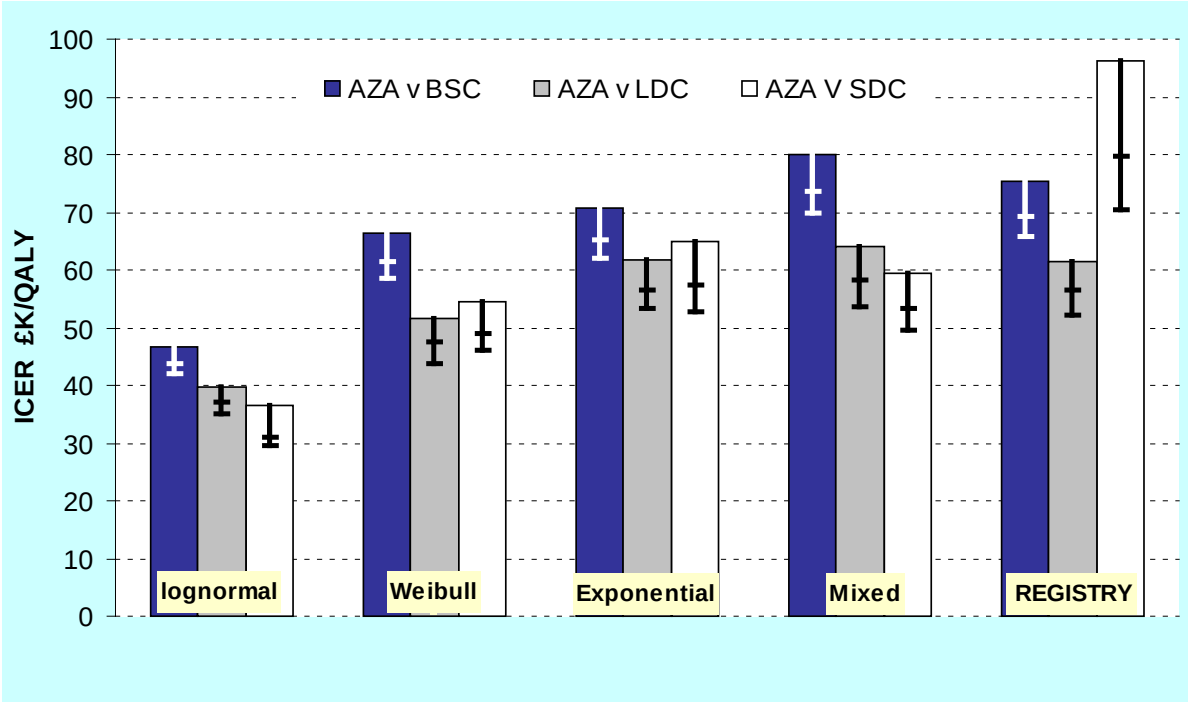
4 Summary and conclusions

In response to the ACD the manufacturer has submitted a new economic analysis which encompasses several fundamental changes from its precursor including: use of updated data to model overall survival (azacitidine subgroups only); the selection of lognormal fits to

overall survival for input for the base case; the use of a different calculation for survival in the AML state. Further modifications concern the handling of adverse events, the use of NHS 2009/10 tariff in costing, the cost of weekend administration of azacitidine, and vial sharing.

Two further CIC modifications to the economic analysis were presented namely a patient access scheme and [redacted]

The manufacturer proposed that their base case analysis, underpinned by lognormal modelling of overall survival, provides the most plausible estimate of the cost effectiveness of azacitidine. However the selection of a lognormal fit is not strongly supported by evidence from the AZA-001 trial or by relevant data from other sources (German registry data and the CALBG 9221 study). The evidence tends to indicate that of the various models that have been explored the Weibull is the most plausible and that the several scenarios explored at NICE's request are also plausible. The ICERs generated by these various models of overall survival submitted by the manufacturer are compared in the diagram below.



ICERs of AZA v comparator for the three pre-selected patient subgroups according to method used in modelling overall survival. The bars in each histogram represent the reduction in ICER consequent on i] vial sharing, ii] vial sharing + PAS.

It is noticeable that the lognormal model delivers substantially lower ICERs than all the other models including the one based on Weibull fits to overall survival. With no vial sharing and without implementation of the PAS no ICER falls below £30,000/QALY, with vial sharing pus the PAS only the lognormal model generates ICERs below £40,000/QALY.

A weakness is a lack of reliable data monitoring time to progression to the AML state. As in previous submissions the RACD modelling has required the assumption, based on clinical opinion, that time spent in AML is the same for all patients; time to progression is then calculated from overall survival minus a single value for time in AML.

The pre-ACD ERGR expressed several concerns regarding the original submission that remain unchanged with respect to the manufacturer's post-ACD submission. These are reiterated below:

- For the comparators, although there is no pooling, the approach taken is to consider the arms of the RCTs included in isolation, effectively breaking randomisation.
- Although the RCT by Fenuax et al (AZA-001) is well conducted it remains open to bias through lack of blinding. There are also concerns about loss to follow-up based on additional information supplied commercially-in-confidence.
- The evidence of different effects in different investigator pre-selected groups is unreliable on the following grounds:
    - Some of the groups, particularly SDC are very small (aza=17; SDC (intensive chemotherapy)=25).
    - The baseline characteristics are often markedly imbalanced, again particularly for the SDC group for the characteristics IPSS classification and karyotype risk; imbalance in the SDC subgroup would be expected to favour AZA.

Caution should be exercised concerning the interpretation of the evidence presented on impact on HRQoL and difference in effect between different investigator pre-selected groups.

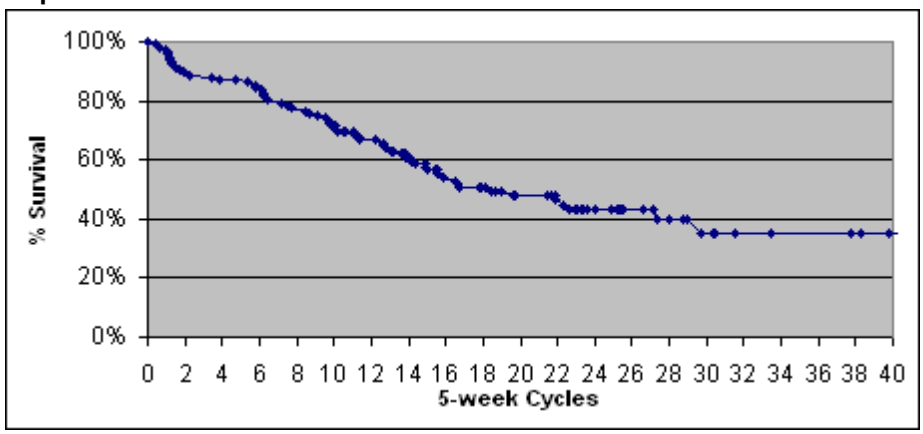
Appendix 1 Mismatch of time unit on axes of survival graphs

The time scale for observed data shown in RACD Fig A1 appeared short compared with what could be expected from the use of "extended" trial data so that the ERG harboured some concern regarding the identity of the fitted curves shown.

The ERG therefore examined the RACD graphs and compared these with those in the model. As an example the BSC-azacitidine subgroup graphs are shown below.

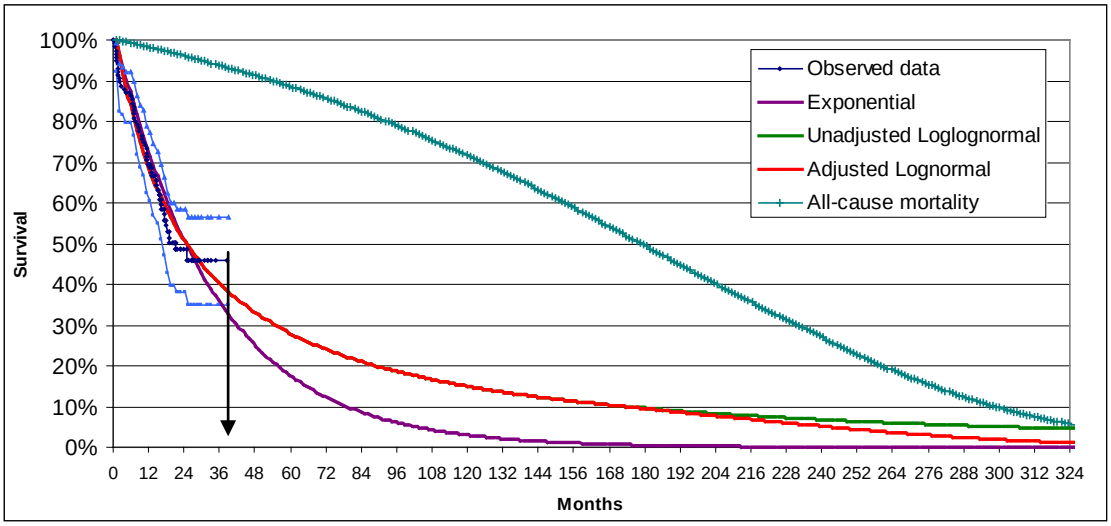
From the model:

Kaplan-Meier survival curve



From the RACD:

Figure A1a: Survival curves for azacitidine (pre-selected for BSC alone)



According to the model graph the observed data extends for 40 5-week cycles = 46 months while according to the RACD graph the observed data extends to ~40 months.

The BSC-alone subgroup graphs are shown below:

From the Model (note the time axis unit in this example is months not 5-week cycles)

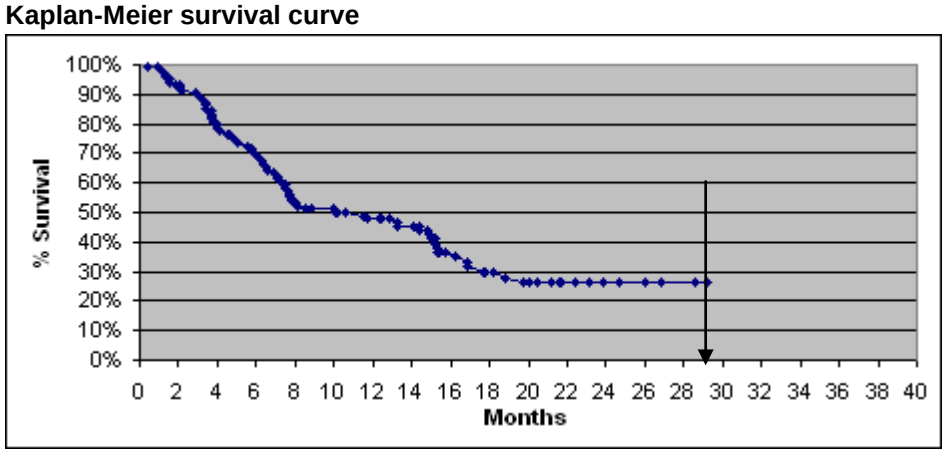
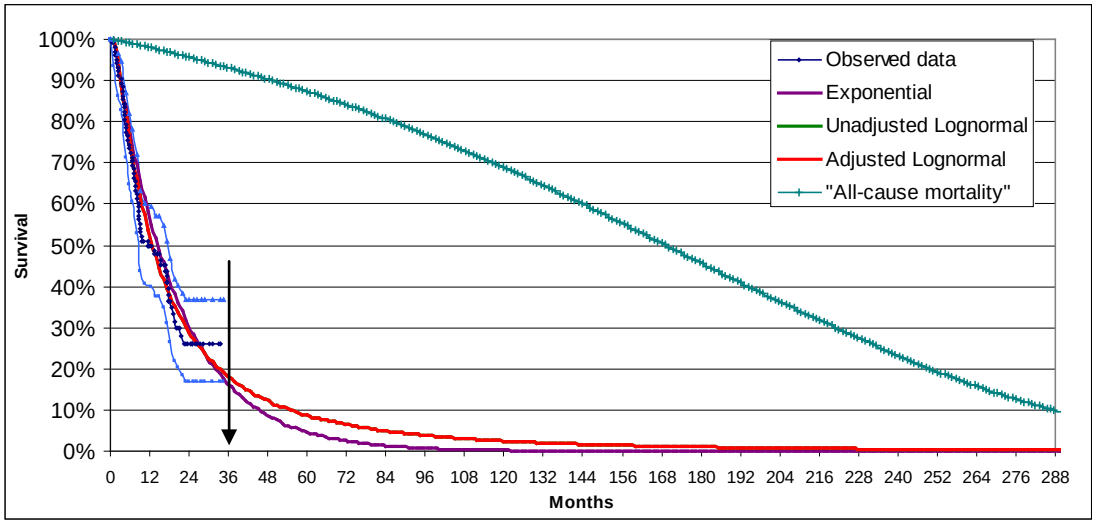


Figure A1d: Survival curves for BSC alone



The model graph implies about 29 months of observation, the RACD graph implies about 36 months.

The RACD states:

In section 3.3 of the ERGR, the ERG also reported that they were unclear as to whether the survival graphs were plotted in months or in 5-week cycles. The survival graphs and the model have been updated so that labelling is consistent and is clear as to which approach has been used.

Unfortunately there is still confusion regarding the time axes. In the model all the graphs (except for the BSC subgroup shown above) have a 5-week cycle as time unit whereas in the RACD document all survival graphs have months as the axis time unit. This makes comparisons / validation between model and submission difficult

Appendix 2 The AZA-001 extension data only applies to azacitidine-treated patients

The RACD states

> After adjusting the economic modelling, the AIC has been recalculated using Study AZA-001 data along with the AZA-001 extension data presented in Celgene's original submission. Based on these new estimates, the exponential is the best fit to the azacitidine (BSC), azacitidine (SDC) and LDC data, and the lognormal is the best fit for the azacitidine (LDC), BSC and SDC data. Further information is provided in the Appendix to this document.

According to the original submission the extension study apparently involves longer follow up of azacitidine-treated patients only, if so the new survival fits will only change for the intervention subgroups.

This is born out when loglogistic fit parameters for the six subgroups are compared between the new model and the originally submitted model. As shown below the control group parameters remain the same across models while for the intervention (azacitidine) subgroups the parameters differ between models.

BSC subgroup

New model loglog		First model loglog	
MLE of λ	**1.496436**	MLE of λ	**1.496436**
MLE of p	**0.093455**	MLE of p	**0.093455**

BSC azacitidine subgroup

New model loglog		First model loglog	
MLE of λ	**1.177896**	MLE of λ	**1.151758**
MLE of p	**0.049372**	MLE of p	**0.048033**

LDC subgroup

New model loglog		First model loglog	
MLE of λ	**1.414978**	MLE of λ	**1.414978**
MLE of p	**0.083889**	MLE of p	**0.083889**

LDC azacitidine subgroup

New model loglog		First model loglog	
MLE of λ	**1.039143**	MLE of λ	**1.083173**
MLE of p	**0.046416**	MLE of p	**0.048282**

SDC subgroup

New model loglog		First model loglog	
MLE of λ	**2.276036**	MLE of λ	**2.276036**
MLE of p	**0.083790**	MLE of p	**0.083790**

SDC azacitidine subgroup

New model loglog		First model loglog	
MLE of λ	**1.110718**	MLE of λ	**1.170957**
MLE of p	**0.053470**	MLE of p	**0.056948**

Appendix 3 Lognormal model parameters are incorrect for the LDC subgroup

The parameters for fits for BSC and LDC subgroups have been copied from the submitted model and pasted below. Note that lognormal and Gompertz parameters are identical for BSC and LDC groups.

BSC

BSC

Weibul parameters

Survival analysis output

Parameter	Coef.
Constant	-3.143877
ln(p)	0.1196

Covariance matrix

Parameter	*Constant*
Constant	0.108624
ln(p)	-0.031482

PSA sampling

Parameter	*Random nos*
Constant	0.33028
ln(p)	0.51614

Estimated parameters

MLE alpha	**1.1270**
MLE Beta	**16.2726**

Log-logistic parameters

Survival analysis output

Parameter	Coef.
Constant	2.370275
ln(gamma)	-0.403086

Covariance matrix

Parameter	*Constant*
Constant	0.014577
ln(gamma)	0.002127

PSA sampling

Parameter	*Random nos*
Constant	0.80251
ln(gamma)	0.18612

Estimated parameters

MLE of λ	**1.49644**
MLE of p	**0.09346**

Exponential

Survival analysis output

Parameter	Coef.
Constant	-2.812007

Estimated parameters

MLE of λ	**0.06008**

LDC

LDC

Weibul parameters

Survival analysis output

Parameter	Coef.
Constant	-3.295527
ln(p)	0.149064

Covariance matrix

Parameter	*Constant*
Constant	0.263285
ln(p)	-0.074147

PSA sampling

Parameter	*Random nos*
Constant	0.48211
ln(p)	0.62809

Estimated parameters

MLE alpha	**1.1607**
MLE Beta	**17.1011**

Log-logistic parameters

Survival analysis output

Parameter	Coef.
Constant	2.478266
ln(gamma)	-0.347114

Covariance matrix

Parameter	*Constant*
Constant	0.034651
ln(gamma)	0.003393

PSA sampling

Parameter	*Random nos*
Constant	0.72309
ln(gamma)	0.32797

Estimated parameters

MLE of λ	**1.41498**
MLE of p	**0.08389**

Exponential

Survival analysis output

Parameter	Coef.
Constant	-2.867637

Estimated parameters

MLE of λ	**0.05683**

BSC

Gompertz *Parameterisation:*

Survival analysis output

Parameter	Coef.
Constant	-2.759109
Gamma	-0.006972

Covariance matrix

Parameter	*Constant*
Constant	0.038961
Gamma	-0.003188

PSA sampling

Parameter	*Random nos*
Constant	0.60451
Gamma	0.86974

Estimated parameters

MLE of β	**-2.75911**
MLE of γ	**-0.00697**

Lognormal *Parameterisation:*

Survival analysis output

Parameter	Coef.
Constant	2.38318
ln(sigma)	0.136185

Covariance matrix

Parameter	*Constant*
Constant	0.015535
ln(sigma)	0.003121

PSA sampling

Parameter	*Random nos*
Constant	0.92058
ln(sigma)	0.56114

Estimated parameters

MLE of β	**2.38318**
MLE of σ	**1.14589**

LDC

Gompertz *Parameterisation:*

Survival analysis output

Parameter	Coef.
Constant	-2.759109
Gamma	-0.006972

Covariance matrix

Parameter	*Constant*
Constant	0.038961
Gamma	-0.003188

PSA sampling

Parameter	*Random nos*
Constant	0.94740
Gamma	0.31089

Estimated parameters

MLE of β	**-2.75911**
MLE of γ	**-0.00697**

Lognormal *Parameterisation:*

Survival analysis output

Parameter	Coef.
Constant	2.38318
ln(sigma)	0.136185

Covariance matrix

Parameter	*Constant*
Constant	0.015535
ln(sigma)	0.003121

PSA sampling

Parameter	*Random nos*
Constant	0.98829
ln(sigma)	0.98224

Estimated parameters

MLE of β	**2.38318**
MLE of σ	**1.14589**

REFERENCES

1 Machin D, Cheung YB, Parmar MKB. Survival analysis; a practical approach. 2 ed. Wiley; 2006.