ESTIMATION OF UTILITIES FOR HEALTH STATES ASSOCIATED WITH DIFFERENT TREATMENTS FOR MYELODYSPLASTIC SYNDROME (MDS)

Report prepared for Pharmion Vidaza

Dr. Donna Rowen, Research Fellow

John Brazier, Professor of Health Economics

School of Health and Related Research (ScHARR), University of Sheffield

July 2008

For further information contact Donna Rowen at:

[redacted]

Telephone number: [redacted]

Fax: [redacted]

E-mail: [redacted]

1. INTRODUCTION

The aim of the project is to determine the preference-based utility value associated with different treatments for Myelodysplastic Syndrome (MDS). No preference-based measures of health-related quality of life were used in trials for MDS. Therefore the utility value associated with different treatments is estimated using non-preference based quality of life measures in existing trial data. The project has three phases. Each of these will be discussed in turn, and the final section provides concluding remarks.

2. PHASE 1

Phase 1 involves the estimation of mean EQ-5D utility values for supportive care and Aza C treatments using publicly available mean quality of life values reported in Kornblith et al. (2002). The EQ-5D score that is attributable to specified EORTC QLQ-C30 functionings and symptoms was calculated using mean values from tables 2 and 3 in Kornblith et al. (2002) and mapping regression coefficients from table 2 in McKenzie and van der Pol (2008). The EQ-5D is the most widely used generic preference-based single index measure of health that can be used to generate QALYs which produces utility scores anchored at 0 for death and 1 for perfect health (Dolan, 1997), where scores below zero indicate that the state is worse than death. The EQ-5D has 5 dimensions: mobility, self-care, usual activity, pain/discomfort and anxiety/depression, each with 3 levels from no problems to extreme problems.

Table 1 shows the EQ-5D score attributable to selected functionings and symptoms of the EORTC QLQ-C30 for the treatments of supportive care and Aza C, across subgroup and time period. The full EQ-5D score has not been calculated as Kornblith et al. (2002) do not report information on mean values for all the functionings and items in the EORTC QLQ-C30. Instead table 1 reports the total EQ-5D score attributable to physical functioning, fatigue, dyspnea, insomnia, social functioning and overall quality of life, which are all functionings and items included in tables 2, 3 and 4 in Kornblith et al. (2002). Although Kornblith et al. (2002) do not report data for all functionings and items included in the EORTC QLQ-C30, arguably data is reported for all functionings and items found to be important in their study of cancer and leukemia.

The subgroups are categorized according to their final quality of life assessment, where subgroup 1 consists mainly of patients assessed once at study entry, subgroup 2 consists mainly of patients assessed twice at baseline and the first follow-up, subgroup 3 consists mainly of patients assessed three times at baseline, first follow-up, second follow-up, and subgroup 4 consists of patients mainly assessed all 4 times. The first assessment for the baseline values was scheduled for study entry, follow-up 1 was scheduled for day 50, follow-up 2 for day 106 and follow-up 3 for day 182. Table 1 also reports the change in EQ-5D score for each follow-up period from the baseline time period and from the previous time period, where the change in EQ-5D score is attributable to physical functioning, fatigue, dyspnea, insomnia, social functioning and overall quality of life.

The results indicate a reduction in EQ-5D score from baseline to the third follow-up of 0.016 using supportive care and a gain of 0.059 using Aza C, which is a difference of 0.075 across the two treatments, where the change in EQ-5D score is attributable only to changes in physical functioning, fatigue, dyspnea, insomnia, social functioning and overall quality of life measured by the EORTC QLQ-C30. The full EQ-5D score or total change in EQ-5D score is not computed here as McKenzie and van der Pol (2008) map the EORTC QLQ-C30 onto the EQ-5D using all functionings and symptoms, but mean values are only available for a subset of functionings and symptoms. The total EQ-5D score would be much higher than the partial EQ-5D score in table 1, at least by 0.2376 in each subgroup for each time period as this is the constant term reported in McKenzie and van der Pol (2008).

3. PHASE 2

Phase 2 involves the estimation of individual level EQ-5D utility values for supportive care and Aza C using the individual level dataset used in Kornblith et al. (2002). The EQ-5D score that is attributable to all EORTC QLQ-C30 functionings, symptoms and global quality of life has been calculated using individual level data in the reconstructed dataset from CALGB that was used in Kornblith et al. (2002) and mapping regression coefficients from table 2 in McKenzie and van der Pol (2008). This phase furthers the analysis in phase 1, but here the full EQ-5D score and total change in EQ-5D score can be computed as McKenzie and van der Pol (2008) map the EORTC QLQ-C30 onto the EQ-5D using all functionings and symptoms, and

these are all available in the recreated individual level dataset originally used in Kornblith et al. (2002). We do not get the same mean age as reported in Kornblith et al. (2002) but get the same *n* overall and the same *n* for each treatment group. This leads us to conclude that the reconstructed individual level dataset may not be identical to the mean level data reported in Kornblith et al. (2002).

The subgroups used in the analysis here differ from those used in Kornblith et al. (2002) as the definitions of the subgroups were not sufficiently precise to enable exact replication.[1] Therefore the subgroups have been defined using the following criteria which is as close to the subgroups used in Kornblith et al. as possible: subgroup 1 consists of patients with one quality of life assessment, subgroup 2 consists of patients with two quality of life assessments, subgroup 3 consists of patients with three quality of life assessments and subgroup 4 consists of patients with four or more quality of life assessments. Quality of life assessments were scheduled at 4 time points as outlined above in phase 1. Patients who crossed over treatments restarted the quality of life assessments at crossover, and the dataset has a maximum of 8 reported assessments.

The analysis is repeated for intention to treat analysis and censored analysis. Intention to treat analysis includes results from all study participants by the groups to which they were allocated at the start of the study, even if they swapped groups or dropped out of the study. Censored analysis removes individuals from the analysis once they crossover treatments. Kornblith et al. state that "After a minimum period of 4 months, those on the supportive care arm could cross over to the Aza C arm based on strict criteria concerning disease progression … patients exited from the supportive care arm within the first 4 months only because of leukemic transformation or platelets less than $20 \times {}^{10^9}\!/_{L}$" (2002: 2442). In the dataset no patients crossed over from Aza C to supportive care and 49 patients crossed over from supportive care to Aza C.

[1] Kornblith et al. state that "Patients were therefore categorized into four subgroups, based on the time of their last quality of life assessment, with subgroups generally coinciding with the number of assessments, as follows: subgroup 1, patients at study entry within 39 days after randomisation, including a few patients with two assessments within this time interval; subgroup 2, mostly consisting of those assessed twice, with the last assessment occurring between days 40 and 82; subgroup 3, mostly consisting of those assessed three times, with the last assessment conducted between days 83 and 159; and subgroup 4, mostly consisting of those assessed four times, with the last assessment conducted between days 160 and 259." (2002: 2442).

Tables 2 and 3 show mean EORTC QLQ-C30 scores for each subgroup for supportive care, defining subgroups using intention to treat and censored analysis respectively. Tables 4 and 5 show mean EORTC QLQ-C30 scores for each subgroup for Aza C, defining subgroups using intention to treat and censored analysis respectively. Tables 2, 3, 4 and 5 replicate the information reported in tables 2 and 3 in Kornblith et al. (2002) with the subgroups as defined above and also include mean scores for all EORTC QLQ-C30 functionings, symptom scales, global quality of life and the overall average score.

Table 6 shows mean EQ-5D score attributable to the EORTC QLQ-C30 across treatment, subgroup and assessment calculated using mapping regression coefficients from table 2 in McKenzie and van der Pol (2008). This indicates that the mean EQ-5D score for patients with 4 or more assessments increases from 0.67 to 0.72 from the first to the fourth assessment for supportive care and from 0.67 to 0.80 for Aza C, which is a difference in gain of 0.08 across the two treatments. The results using intention to treat and censored analysis are almost identical up to the fourth assessment, but differ for assessments 5 onwards for supportive care treatment. Figures 1 and 4 are error bar plots of EQ-5D score showing means and 95% confidence intervals by treatment and assessment using intention to treat and censored analysis respectively.

Figures 2 and 5 are error bar plots of EQ-5D score by subgroup for those receiving supportive care using intention to treat and censored analysis respectively. They indicate that EQ-5D score falls after the first assessment for subgroups 2 and 3, suggesting that the treatment may initially have a negative effect. Subgroup 4 has an initial increase, followed by a reduction and subsequently an increase in the fourth assessment that outweighs any reduction and results in an overall increased mean EQ-5D score. Figures 3 and 6 are error bar plots of EQ-5D score by subgroup for those receiving Aza C using intention to treat and censored analysis respectively. These plots indicate a similar pattern for Aza C in comparison to supportive care where EQ-5D score falls after the first assessment for subgroups 2 and 3, but for subgroup 3 this is followed by a positive effect that outweighs the initial reduction. Subgroup 4 has a increase in EQ-5D score in each assessment.

4. PHASE 3

Phase 3 involves the estimation of individual level SF-6D and EQ-5D utility values for intensive induction chemotherapy (IC) and non-intensive chemotherapy/best supportive care (NIC) using the individual level dataset used in Sekeres et al. (2004).

Table 7 shows SF-6D scores by treatment and time period. The SF-6D is a generic preference-based single index measure of health that can be used to generate QALYs and hence which can be used in cost-utility analysis. The SF-6D is a utility value that is anchored at 1 for full health and 0 for dead, hence a reduction in utility score represents a reduction in health-related quality of life. The SF-6D has 6 dimensions: physical functioning, role limitations, social functioning, pain, mental health, vitality, each with between 4 and 6 levels depending on the dimensions. The SF-6D utility score produced here has been derived from the SF-12 data in the dataset used in Sekeres et al. (2004). The SF-6D utility score has been produced using an algorithm that weights the domain scores using Brazier and Roberts (2004).

Table 8 shows EQ-5D scores by treatment and time period calculated using the Franks et al. (2004) model to predict EQ-5D utility scores. Franks et al. regress the EQ-5D utility score on PCS-12 and MCS-12, squared terms and cross-products using an ordinary least squares model (OLS). PCS and MCS are the physical and mental component summary scores estimated using factor analysis and shown to contain most of the information contained in the 8 dimensions of the SF-36 (Ware et al., 1995). In accordance with this approach PCS-12 and MCS-12 are centred on the means used by Franks et al. and the published coefficients are used to produce predicted EQ-5D utility scores.[2]

Table 9 shows EQ-5D scores by treatment and time period calculated using the Gray et al. (2006) algorithm to predict EQ-5D utility scores. The Gray et al. approach was developed as an improvement to already existing mapping approaches including the approach by Franks et al. (2004) and uses a more sophisticated mapping approach

[2] Franks et al. (2004) estimate other models but these are not analysed here as these models use demographic variables not available in the dataset used here. Furthermore Franks et al. found that more complex models explained only minimally additional variance.

than previous approaches. Gray et al. use a response mapping approach that uses a multinomial logit model to estimate the probability that a respondent will choose a particular level for each dimension of the EQ-5D using responses to the 12 items included in the SF-12 (general health, climbing stairs, moderate activities, accomplish less due to physical health, work limitations, accomplish less due to emotional problems, work carefully, pain interference, calm, energy, down-hearted and low, interference with social activities). Subsequently predicted EQ-5D level responses for each dimension are generated using Monte Carlo simulation methods and the corresponding EQ-5D utility score for that health state is calculated.

Tables 7, 8 and 9 show that the mean utility scores for IC increase from baseline to 14 weeks: from 0.66 to 0.72 for SF-6D score; from 0.59 to 0.66 for EQ-5D score estimated using the Franks et al. approach; and from 0.62 to 0.71 for EQ-5D score estimated using the Gray et al. approach. Mean utility scores for NIC increase from baseline to 14 weeks: from 0.67 to 0.70 for SF-6D score; from 0.62 to 0.66 for EQ-5D score estimated using the Franks et al. approach; and from 0.63 to 0.76 for EQ-5D score estimated using the Gray et al. approach. This represents a difference in gain across treatments of 0.03 for the SF-6D score, 0.03 for EQ-5D score estimated using the Franks et al. approach, and 0.04 for EQ-5D score estimated using the Gray et al. approach.

Figure 7 is an error bar plot showing mean SF-6D score by treatment and time period using the individual level dataset used in Sekeres et al. (2004). Figures 8 and 9 are error bar plots showing mean EQ-5D score estimated using the Franks et al. (2004) and Gray et al. (2006) approaches respectively. Utility scores calculated using the 3 different methods are generally consistent. All show an improvement in mean utility score from baseline to 14 weeks for both treatments, yet the Gray et al. approach shows a higher change in utility score for NIC and the other approaches show a higher change for IC. Utility scores for IC for all methods are reduced at 2 weeks in comparison to baseline and then increase in following assessments, with the exception of the assessment at 6 weeks using the Gray et al. (2006) approach. The pattern of utility scores for NIC varies according to the method used to calculate the utility scores.

5. CONCLUSION

Preference-based utility values associated with different treatments for Myelodysplastic Syndrome (MDS) have been generated using published mean values for a subset of EORTC QLQ-C30 functionings and symptoms in phase one, using individual level EORTC QLQ-C30 functionings and symptoms in phase 2, and using individual level SF-36 scores in phase 3. Table 10 summarises the change in utility score estimated for each treatment in each phase of the analysis for subgroups with 4 assessments in phases 1 and 2 and 5 assessments in phase 3. Results in phases 2 and 3 indicate that all treatments analysed increase utility scores after a considerable time following treatment, yet utility was not always improved in the first assessment after baseline. Phase 1 has inconsistent results for supportive care in comparison to phases 2 and 3, which is most likely due to the limitations of the published mean values used in the analysis.

References

Brazier J. E., Roberts J. R., 2004. The estimation of a preference-based index from the SF-12. *Medical Care*, 42: 851-859.

Dolan, P., 1997. Modeling Valuations for EuroQol Health States. *Medical Care* 35:1095-1108.

Franks, P., et al., 2004. Mapping the SF-12 to the EuroQol EQ-5D Index in a National US Sample. *Medical Decision Making* 24: 247-254.

Gray, A. M., Rivero-Arias, O., Clarke, P. M., 2006. Estimating the Association between SF-12 Responses and EQ-5D Utility Values by Response Mapping. *Medical Decision Making* 26: 18-29.

McKenzie, L., and van der Pol, M., 2008. Mapping the EORTC QLQ C-30 onto the EQ-5D instrument: the potential to estimate QALYs without generic preference data. HERC, University of Aberdeen.

Kornblith, A. B. et al., 2002. Impact of Azacytidine on the Quality of Life of Patients With Myelodysplastic Syndrome Treated in a Randomized Phase III Trial: A Cancer and Leukemia Group B Study, *Journal of Clinical Oncology*, 20: 2441-2452.

Sekeres, M. A. et al., 2004. Decision-making and quality of life in older adults with acute myeloid leukemia or advanced myelodysplastic syndrome, *Leukemia*, 18: 809-816.

APPENDIX

Table 1 EQ-5D score attributable to selected functionings and symptoms of the EORTC QLQ-C30 across treatment, subgroup and time period

		Supportive care				*Aza C*			
	Subgroup	*Baseline*	*Follow up 1*	*Follow up 2*	*Follow up 3*	*Baseline*	*Follow up 1*	*Follow up 2*	*Follow up 3*
Physical functioning	SG 1	0.021				0.021			
	SG 2	0.031	0.022			0.028	0.024		
	SG 3	0.025	0.025	0.020		0.022	0.021	0.027	
	SG 4	0.028	0.027	0.027	0.026	0.026	0.026	0.028	0.031
Fatigue	SG 1	-0.100				-0.100			
	SG 2	-0.090	-0.100			-0.080	-0.090		
	SG 3	-0.072	-0.089	-0.099		-0.093	-0.082	-0.081	
	SG 4	-0.083	-0.080	-0.080	-0.089	-0.083	-0.078	-0.066	-0.055
Dyspnea	SG 1	0.014				0.016			
	SG 2	0.015	0.017			0.012	0.012		
	SG 3	0.011	0.010	0.013		0.017	0.017	0.013	
	SG 4	0.010	0.012	0.014	0.012	0.014	0.013	0.011	0.009
Insomnia	SG 1	0.001				0.001			
	SG 2	0.001	0.001			0.001	0.001		
	SG 3	0.001	0.001	0.001		0.001	0.001	0.001	
	SG 4	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Social functioning	SG 1	0.013				0.012			
	SG 2	0.012	0.009			0.015	0.011		
	SG 3	0.014	0.015	0.012		0.016	0.014	0.014	
	SG 4	0.015	0.016	0.015	0.014	0.014	0.014	0.014	0.015
Overall QoL	SG 1	0.074				0.075			
	SG 2	0.086	0.061			0.095	0.080		
	SG 3	0.081	0.087	0.063		0.084	0.087	0.101	
	SG 4	0.091	0.094	0.091	0.082	0.086	0.091	0.105	0.117
Total EQ-5D score attributable to above functionings and items	SG 1	0.022				0.026			
	SG 2	0.055	0.011			0.070	0.037		
	SG 3	0.061	0.049	0.010		0.048	0.057	0.075	
	SG 4	0.063	0.071	0.068	0.047	0.059	0.067	0.095	0.118
Change in EQ-5D score from baseline time period	SG 1								
	SG 2		-0.043				-0.033		
	SG 3		-0.012	-0.051			0.009	0.027	
	SG 4		0.008	0.005	-0.016		0.008	0.035	0.059
Change in EQ-5D score from previous time period	SG 1								
	SG 2		-0.043				-0.033		
	SG 3		-0.012	-0.039			0.009	0.018	
	SG 4		0.008	-0.003	-0.021		0.008	0.027	0.023

Table 2 Means of EORTC QLQ-C30 for each subgroup for supportive care, defining subgroups using intention to treat analysis

		Assessment											
	Subgroup	*1*			*2*			*3*			*4*		
		N	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*
Average of responses to all EORTC QLQ-C30 questions	1	17	34.40	14.26		.	.		.	.		.	.
	2	9	26.05	13.02	9	38.23	20.85		.	.		.	.
	3	22	23.81	15.69	22	24.43	16.46	22	29.84	18.66		.	.
	4	44	25.35	14.23	44	23.98	12.79	44	25.79	15.89	44	23.81	15.74
Physical functioning	1	17	53	28		.	.		.	.		.	.
	2	9	78	19	9	56	34		.	.		.	.
	3	22	65	29	22	67	32	22	54	35		.	.
	4	44	70	24	44	69	22	44	67	24	44	69	27
Role functioning	1	17	50	45		.	.		.	.		.	.
	2	9	72	36	9	39	42		.	.		.	.
	3	22	73	37	22	64	38	22	59	37		.	.
	4	44	69	34	44	70	38	44	68	37	44	73	35
Dyspnea	1	17	35	28		.	.		.	.		.	.
	2	9	37	11	9	44	29		.	.		.	.
	3	22	26	23	22	26	20	22	33	25		.	.
	4	44	26	20	44	30	20	44	32	19	44	26	23
Pain	1	17	19.66	28.53		.	.		.	.		.	.
	2	9	18.56	13.77	9	23.94	25.07		.	.		.	.
	3	22	10.50	19.42	22	16.52	25.52	22	15.07	26.57		.	.
	4	44	15.03	22.56	44	16.50	20.13	44	15.41	24.61	44	16.14	23.77
Fatigue	1	17	46.85	22.02		.	.		.	.		.	.
	2	9	42.78	18.61	9	49.00	23.60		.	.		.	.
	3	22	33.60	27.24	22	38.30	25.34	22	43.09	28.14		.	.
	4	44	39.32	24.02	44	37.38	18.97	44	39.53	20.26	44	37.64	21.86
Insomnia/Sleep disturbance	1	17	27	30		.	.		.	.		.	.

		Assessment											
	Subgroup	*1*			*2*			*3*			*4*		
		N	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*
	2	9	18	17	9	37	42		.	.		.	.
	3	22	24	23	22	23	21	22	26	25		.	.
	4	44	23	25	44	23	25	44	23	25	44	17	21
Nausea/Vomiting	1	17	10.31	15.80		.	.		.	.		.	.
	2	9	1.83	5.50	9	7.33	8.70		.	.		.	.
	3	22	3.00	9.71	22	3.00	8.27	22	5.27	15.69		.	.
	4	44	3.00	7.36	44	2.25	7.62	44	4.13	12.39	44	7.51	15.33
Constipation	1	17	8	26		.	.		.	.		.	.
	2	9	15	24	9	11	23		.	.		.	.
	3	22	18	32	22	12	24	22	8	23		.	.
	4	44	8	18	44	8	18	44	11	21	44	13	22
Diarrhea	1	17	8	15		.	.		.	.		.	.
	2	9	11	23	9	11	17		.	.		.	.
	3	22	5	12	22	5	12	22	6	17		.	.
	4	44	5	11	44	2	8	44	4	11	44	3	10
Cognitive functioning	1	17	79.03	16.44		.	.		.	.		.	.
	2	9	77.44	16.85	9	75.67	22.38		.	.		.	.
	3	22	83.84	19.09	22	84.57	16.46	22	83.07	19.47		.	.
	4	44	82.74	19.02	44	84.99	17.15	44	81.97	22.41	44	85.41	18.00
Emotional functioning	1	17	65.25	23.36		.	.		.	.		.	.
	2	9	68.11	16.55	9	65.33	20.24		.	.		.	.
	3	22	79.47	15.73	22	80.74	15.21	22	78.80	15.98		.	.
	4	44	68.20	18.23	44	72.85	15.16	44	73.27	18.24	44	77.48	16.76
Financial impact	1	17	27	32		.	.		.	.		.	.
	2	9	7	15	9	22	23		.	.		.	.
	3	22	21	33	22	26	31	22	20	22		.	.
	4	44	19	32	44	18	31	44	19	30	44	21	34
Overall quality of life	1	17	46.53	14.48		.	.		.	.		.	.

		Assessment											
	Subgroup	*1*			*2*			*3*			*4*		
		N	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*
	2	9	53.56	25.77	9	37.00	23.65		.	.		.	.
	3	22	53.48	27.53	22	56.11	27.68	22	41.34	27.14		.	.
	4	44	55.83	21.22	44	58.77	18.54	44	54.67	20.45	44	60.19	19.43
Appetite loss	1	17	25	35		.	.		.	.		.	.
	2	9	15	24	9	37	39		.	.		.	.
	3	22	12	19	22	11	21	22	11	16		.	.
	4	44	20	24	44	16	21	44	15	25	44	15	22
Social functioning	1	17	64.38	17.63		.	.		.	.		.	.
	2	9	60.83	26.48	9	38.61	34.18		.	.		.	.
	3	22	73.95	27.68	22	73.20	27.19	22	61.00	29.70		.	.
	4	44	75.45	23.82	44	78.56	25.88	44	72.03	26.28	44	72.05	25.01

Table 3 Means of EORTC QLQ-C30 for each subgroup for supportive care, defining subgroups using censored analysis

		Assessment															
	Subgroup	*1*			*2*			*3*			*4*			*5*			*6*	
		N	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*					
Average of responses to all EORTC QLQ-C30 questions	1	17	34.40	14.26		.	.		.	.		.	.		.	.		.
	2	9	26.05	13.02	9	38.23	20.85		.	.		.	.		.	.		.
	3	22	23.81	15.69	21	24.81	16.77	20	30.61	18.87		.	.		.	.		.
	4	44	25.35	14.23	43	24.05	12.94	39	26.48	16.55	14	23.17	13.79	4	22.65	22.57	1	54.80	.
Physical functioning	1	17	53	28		.	.		.	.		.	.		.	.		.	.
	2	9	78	19	9	56	34		.	.		.	.		.	.		.	.
	3	22	65	29	21	66	32	20	52	35		.	.		.	.		.	.
	4	44	70	24	43	68	22	39	65	25	14	69	22	4	70	38	1	20	.
Role functioning	1	17	50	45		.	.		.	.		.	.		.	.		.	.
	2	9	72	36	9	39	42		.	.		.	.		.	.		.	.
	3	22	73	37	21	62	38	20	58	37		.	.		.	.		.	.
	4	44	69	34	43	71	38	39	68	39	14	79	32	4	75	50	1	0	.
Dyspnea	1	17	35	28		.	.		.	.		.	.		.	.		.	.
	2	9	37	11	9	44	29		.	.		.	.		.	.		.	.
	3	22	26	23	21	25	21	20	35	25		.	.		.	.		.	.
	4	44	26	20	43	31	20	39	34	19	14	31	24	4	25	32	1	66	.
Pain	1	17	19.66	28.53		.	.		.	.		.	.		.	.		.	.
	2	9	18.56	13.77	9	23.94	25.07		.	.		.	.		.	.		.	.
	3	22	10.50	19.42	21	17.31	25.88	20	12.43	22.76		.	.		.	.		.	.
	4	44	15.03	22.56	43	16.12	20.20	39	15.69	25.78	14	15.36	25.55	4	4.13	8.25	1	.00	.
Fatigue	1	17	46.85	22.02		.	.		.	.		.	.		.	.		.	.
	2	9	42.78	18.61	9	49.00	23.60		.	.		.	.		.	.		.	.
	3	22	33.60	27.24	21	40.22	24.38	20	45.75	27.65		.	.		.	.		.	.

		Assessment																
	Subgroup	*1*			*2*			*3*			*4*			*5*			*6*	
		N	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*					
	4	44	39.32	24.02	43	38.27	18.27	39	40.65	20.15	14	38.55	22.79	4	35.83	29.09	1	88.67	.
Insomnia/Sleep disturbance	1	17	27	30		.	.		.	.		.	.		.	.		.	.
	2	9	18	17	9	37	42		.	.		.	.		.	.		.	.
	3	22	24	23	21	24	21	20	28	25		.	.		.	.		.	.
	4	44	23	25	43	24	25	39	24	25	14	19	21	4	25	17	1	33	.
Nausea/Vomiting	1	17	10.31	15.80		.	.		.	.		.	.		.	.		.	.
	2	9	1.83	5.50	9	7.33	8.70		.	.		.	.		.	.		.	.
	3	22	3.00	9.71	21	3.14	8.44	20	5.80	16.40		.	.		.	.		.	.
	4	44	3.00	7.36	43	1.92	7.39	39	2.96	8.36	14	1.18	4.41	4	8.25	16.50	1	.00	.
Constipation	1	17	8	26		.	.		.	.		.	.		.	.		.	.
	2	9	15	24	9	11	23		.	.		.	.		.	.		.	.
	3	22	18	32	21	11	24	20	3	10		.	.		.	.		.	.
	4	44	8	18	43	7	15	39	8	20	14	5	12	4	8	17	1	33	.
Diarrhea	1	17	8	15		.	.		.	.		.	.		.	.		.	.
	2	9	11	23	9	11	17		.	.		.	.		.	.		.	.
	3	22	5	12	21	5	12	20	7	17		.	.		.	.		.	.
	4	44	5	11	43	2	9	39	4	11	14	5	12	4	8	17	1	0	.
Cognitive functioning	1	17	79.03	16.44		.	.		.	.		.	.		.	.		.	.
	2	9	77.44	16.85	9	75.67	22.38		.	.		.	.		.	.		.	.
	3	22	83.84	19.09	21	85.45	16.33	20	81.38	19.64		.	.		.	.		.	.
	4	44	82.74	19.02	43	85.03	17.35	39	80.53	23.25	14	83.14	18.65	4	91.50	17.00	1	66.00	.
Emotional functioning	1	17	65.25	23.36		.	.		.	.		.	.		.	.		.	.
	2	9	68.11	16.55	9	65.33	20.24		.	.		.	.		.	.		.	.
	3	22	79.47	15.73	21	81.44	15.22	20	77.95	16.49		.	.		.	.		.	.
	4	44	68.20	18.23	43	72.62	15.26	39	72.02	18.53	14	75.75	14.81	4	74.56	15.43	1	49.50	.
Financial impact	1	17	27	32		.	.		.	.		.	.		.	.		.	.

		Assessment																
	Subgroup	*1*			*2*			*3*			*4*			*5*			*6*	
		N	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*					
	2	9	7	15	9	22	23		.	.		.	.		.	.		.	.
	3	22	21	33	21	24	30	20	18	23		.	.		.	.		.	.
	4	44	19	32	43	18	31	39	19	28	14	17	28	4	8	17	1	33	.
Overall quality of life	1	17	46.53	14.48		.	.		.	.		.	.		.	.		.	.
	2	9	53.56	25.77	9	37.00	23.65		.	.		.	.		.	.		.	.
	3	22	53.48	27.53	21	56.40	28.32	20	39.63	27.80		.	.		.	.		.	.
	4	44	55.83	21.22	43	58.71	18.76	39	54.21	21.03	14	56.50	21.70	4	58.38	15.34	1	25.00	.
Appetite loss	1	17	25	35		.	.		.	.		.	.		.	.		.	.
	2	9	15	24	9	37	39		.	.		.	.		.	.		.	.
	3	22	12	19	21	11	22	20	10	16		.	.		.	.		.	.
	4	44	20	24	43	15	21	39	15	25	14	9	15	4	25	32	1	33	.
Social functioning	1	17	64.38	17.63		.	.		.	.		.	.		.	.		.	.
	2	9	60.83	26.48	9	38.61	34.18		.	.		.	.		.	.		.	.
	3	22	73.95	27.68	21	73.55	27.81	20	61.33	31.09		.	.		.	.		.	.
	4	44	75.45	23.82	43	79.62	25.21	39	72.78	26.79	14	75.86	22.53	4	74.75	32.12	1	16.50	.

Table 4 Means of EORTC QLQ-C30 for each subgroup for Aza C, defining subgroups using intention to treat analysis

		Assessment											
	Subgroup	*1*			*2*			*3*			*4*		
		N	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*
Average of responses to all EORTC QLQ-C30 questions	1	14	35.88	13.95		.	.		.	.		.	.
	2	11	23.62	14.34	11	30.78	15.53		.	.		.	.
	3	18	31.51	13.27	18	33.51	15.19	18	28.65	18.24		.	.
	4	56	27.56	16.49	56	26.26	14.11	56	22.58	14.02	56	18.19	14.29
Physical functioning	1	14	53	23		.	.		.	.		.	.
	2	11	67	26	11	59	31		.	.		.	.
	3	18	53	21	18	51	22	18	63	31		.	.
	4	56	64	27	56	65	27	56	70	26	56	78	23
Role functioning	1	14	43	43		.	.		.	.		.	.
	2	11	68	34	11	55	42		.	.		.	.
	3	18	47	37	18	50	34	18	62	33		.	.
	4	56	58	37	56	66	36	56	68	35	56	73	37
Dyspnea	1	14	40	23		.	.		.	.		.	.
	2	11	27	13	11	30	18		.	.		.	.
	3	18	46	31	18	40	24	18	33	20		.	.
	4	56	36	24	56	34	21	56	28	23	56	21	23
Pain	1	14	30.50	26.04		.	.		.	.		.	.
	2	11	10.50	15.25	11	15.00	26.04		.	.		.	.
	3	18	9.19	19.89	18	11.92	20.25	18	7.82	25.06		.	.
	4	56	13.26	20.25	56	11.50	17.24	56	9.74	18.10	56	7.96	14.42
Fatigue	1	14	47.21	22.25		.	.		.	.		.	.
	2	11	36.00	17.11	11	40.00	15.77		.	.		.	.
	3	18	46.57	24.79	18	44.04	19.69	18	39.17	20.87		.	.
	4	56	39.56	24.13	56	36.37	19.40	56	31.63	18.17	56	26.16	21.03
Insomnia/Sleep disturbance	1	14	35	27		.	.		.	.		.	.
	2	11	21	31	11	30	23		.	.		.	.

		Assessment											
	Subgroup	*1*			*2*			*3*			*4*		
		N	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*
	3	18	24	25	18	24	25	18	26	33		.	.
	4	56	33	33	56	25	28	56	18	23	56	16	24
Nausea/Vomiting	1	14	10.64	15.42		.	.		.	.		.	.
	2	11	1.50	4.97	11	12.00	16.65		.	.		.	.
	3	18	6.42	11.51	18	11.92	18.60	18	11.00	16.01		.	.
	4	56	7.08	16.02	56	9.14	17.81	56	7.66	12.96	56	5.01	9.91
Constipation	1	14	18	35		.	.		.	.		.	.
	2	11	12	22	11	6	20		.	.		.	.
	3	18	11	16	18	24	27	18	29	34		.	.
	4	56	12	20	56	17	24	56	18	23	56	17	24
Diarrhea	1	14	7	14		.	.		.	.		.	.
	2	11	0	0	11	18	27		.	.		.	.
	3	18	9	15	18	7	18	18	6	13		.	.
	4	56	4	12	56	7	17	56	4	12	56	5	16
Cognitive functioning	1	14	68.64	19.55		.	.		.	.		.	.
	2	11	87.68	15.25	11	84.59	17.66		.	.		.	.
	3	18	81.17	17.22	18	82.08	14.77	18	81.19	20.75		.	.
	4	56	84.29	18.67	56	85.17	15.13	56	89.12	13.83	56	88.80	16.51
Emotional functioning	1	14	71.35	17.42		.	.		.	.		.	.
	2	11	76.18	21.13	11	73.80	13.84		.	.		.	.
	3	18	68.11	21.60	18	70.42	20.18	18	71.78	16.26		.	.
	4	56	74.79	20.57	56	75.94	15.51	56	79.30	17.64	56	83.96	15.64
Financial impact	1	14	17	21		.	.		.	.		.	.
	2	11	18	23	11	27	39		.	.		.	.
	3	18	20	28	18	33	32	18	30	38		.	.
	4	56	22	28	56	25	32	56	28	31	56	25	30
Overall quality of life	1	14	47.50	14.88		.	.		.	.		.	.
	2	11	58.95	20.27	11	52.95	18.32		.	.		.	.

		Assessment											
	Subgroup	*1*			*2*			*3*			*4*		
		N	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*
	3	18	50.39	25.80	18	48.53	22.62	18	60.72	22.46		.	.
	4	56	54.00	22.24	56	58.00	19.57	56	65.66	18.20	56	73.59	19.08
Appetite loss	1	14	24	36		.	.		.	.		.	.
	2	11	24	26	11	36	35		.	.		.	.
	3	18	20	28	18	31	35	18	29	30		.	.
	4	56	20	27	56	17	22	56	13	22	56	12	22
Social functioning	1	14	60.32	25.87		.	.		.	.		.	.
	2	11	75.50	27.37	11	58.77	34.42		.	.		.	.
	3	18	77.47	21.51	18	65.39	28.93	18	69.08	25.85		.	.
	4	56	70.49	25.56	56	69.02	24.90	56	71.65	22.19	56	77.34	21.56

Table 5 Means of EORTC QLQ-C30 for each subgroup for Aza C, defining subgroups using censored analysis

		Assessment											
	Subgroup	*1*			*2*			*3*			*4*		
		N	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*
Average of responses to all EORTC QLQ-C30 questions	1	14	35.88	13.95		.	.		.	.		.	.
	2	11	23.62	14.34	11	30.78	15.53		.	.		.	.
	3	18	31.51	13.27	18	33.51	15.19	18	28.65	18.24		.	.
	4	56	27.56	16.49	56	26.26	14.11	56	22.58	14.02	56	18.19	14.29
Physical functioning	1	14	53	23		.	.		.	.		.	.
	2	11	67	26	11	59	31		.	.		.	.
	3	18	53	21	18	51	22	18	63	31		.	.
	4	56	64	27	56	65	27	56	70	26	56	78	23
Role functioning	1	14	43	43		.	.		.	.		.	.
	2	11	68	34	11	55	42		.	.		.	.
	3	18	47	37	18	50	34	18	62	33		.	.
	4	56	58	37	56	66	36	56	68	35	56	73	37
Dyspnea	1	14	40	23		.	.		.	.		.	.
	2	11	27	13	11	30	18		.	.		.	.
	3	18	46	31	18	40	24	18	33	20		.	.
	4	56	36	24	56	34	21	56	28	23	56	21	23
Pain	1	14	30.50	26.04		.	.		.	.		.	.
	2	11	10.50	15.25	11	15.00	26.04		.	.		.	.
	3	18	9.19	19.89	18	11.92	20.25	18	7.82	25.06		.	.
	4	56	13.26	20.25	56	11.50	17.24	56	9.74	18.10	56	7.96	14.42
Fatigue	1	14	47.21	22.25		.	.		.	.		.	.
	2	11	36.00	17.11	11	40.00	15.77		.	.		.	.
	3	18	46.57	24.79	18	44.04	19.69	18	39.17	20.87		.	.
	4	56	39.56	24.13	56	36.37	19.40	56	31.63	18.17	56	26.16	21.03
Insomnia/Sleep disturbance	1	14	35	27		.	.		.	.		.	.
	2	11	21	31	11	30	23		.	.		.	.

		Assessment											
	Subgroup	1			2			3			4		
		N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
	3	18	24	25	18	24	25	18	26	33		.	.
	4	56	33	33	56	25	28	56	18	23	56	16	24
Nausea/Vomiting	1	14	10.64	15.42		.	.		.	.		.	.
	2	11	1.50	4.97	11	12.00	16.65		.	.		.	.
	3	18	6.42	11.51	18	11.92	18.60	18	11.00	16.01		.	.
	4	56	7.08	16.02	56	9.14	17.81	56	7.66	12.96	56	5.01	9.91
Constipation	1	14	18	35		.	.		.	.		.	.
	2	11	12	22	11	6	20		.	.		.	.
	3	18	11	16	18	24	27	18	29	34		.	.
	4	56	12	20	56	17	24	56	18	23	56	17	24
Diarrhea	1	14	7	14		.	.		.	.		.	.
	2	11	0	0	11	18	27		.	.		.	.
	3	18	9	15	18	7	18	18	6	13		.	.
	4	56	4	12	56	7	17	56	4	12	56	5	16
Cognitive functioning	1	14	68.64	19.55		.	.		.	.		.	.
	2	11	87.68	15.25	11	84.59	17.66		.	.		.	.
	3	18	81.17	17.22	18	82.08	14.77	18	81.19	20.75		.	.
	4	56	84.29	18.67	56	85.17	15.13	56	89.12	13.83	56	88.80	16.51
Emotional functioning	1	14	71.35	17.42		.	.		.	.		.	.
	2	11	76.18	21.13	11	73.80	13.84		.	.		.	.
	3	18	68.11	21.60	18	70.42	20.18	18	71.78	16.26		.	.
	4	56	74.79	20.57	56	75.94	15.51	56	79.30	17.64	56	83.96	15.64
Financial impact	1	14	17	21		.	.		.	.		.	.
	2	11	18	23	11	27	39		.	.		.	.
	3	18	20	28	18	33	32	18	30	38		.	.
	4	56	22	28	56	25	32	56	28	31	56	25	30
Overall quality of life	1	14	47.50	14.88		.	.		.	.		.	.
	2	11	58.95	20.27	11	52.95	18.32		.	.		.	.

		Assessment											
	Subgroup	*1*			*2*			*3*			*4*		
		N	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*
	3	18	50.39	25.80	18	48.53	22.62	18	60.72	22.46		.	.
	4	56	54.00	22.24	56	58.00	19.57	56	65.66	18.20	56	73.59	19.08
Appetite loss	1	14	24	36		.	.		.	.		.	.
	2	11	24	26	11	36	35		.	.		.	.
	3	18	20	28	18	31	35	18	29	30		.	.
	4	56	20	27	56	17	22	56	13	22	56	12	22
Social functioning	1	14	60.32	25.87		.	.		.	.		.	.
	2	11	75.50	27.37	11	58.77	34.42		.	.		.	.
	3	18	77.47	21.51	18	65.39	28.93	18	69.08	25.85		.	.
	4	56	70.49	25.56	56	69.02	24.90	56	71.65	22.19	56	77.34	21.56

Table 6 EQ-5D score attributable to the EORTC QLQ-C30 across treatment, subgroup and assessment

	Assessment																							
	1			2			3			4			5			6			7			8		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
Supportive care, ITT analysis																								
1	17	.56	.22		.	.		.	.		.	.		.	.		.	.		.	.		.	.
2	9	.70	.18	9	.51	.28		.	.		.	.		.	.		.	.		.	.		.	.
3	22	.72	.23	22	.68	.24	22	.64	.25		.	.		.	.		.	.		.	.		.	.
4	44	.67	.22	44	.69	.20	44	.68	.24	44	.72	.22	34	.78	.19	24	.80	.23	4	.81	.29	1	.89	.
Supportive care, censored analysis																								
1	17	.56	.22		.	.		.	.		.	.		.	.		.	.						
2	9	.70	.18	9	.51	.28		.	.		.	.		.	.		.	.						
3	22	.72	.23	21	.67	.25	20	.63	.26		.	.		.	.		.	.						
4	44	.67	.22	43	.69	.20	39	.67	.25	14	.72	.22	4	.75	.27	1	.32	.						
Aza C, ITT analysis																								
1	14	.52	.24		.	.		.	.		.	.												
2	11	.72	.21	11	.64	.23		.	.		.	.												
3	18	.62	.20	18	.61	.21	18	.72	.17		.	.												
4	56	.67	.23	56	.70	.20	56	.74	.20	56	.80	.21												
Aza C, censored analysis																								
1	14	.52	.24		.	.		.	.		.	.												
2	11	.72	.21	11	.64	.23		.	.		.	.												
3	18	.62	.20	18	.61	.21	18	.72	.17		.	.												
4	56	.67	.23	56	.70	.20	56	.74	.20	56	.80	.21												

Figure 1 Error bar plot of EQ-5D score by treatment and assessment using intention to treat analysis

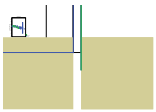

Figure 2 Error bar plot of EQ-5D score by subgroup for those receiving supportive care using intention to treat analysis

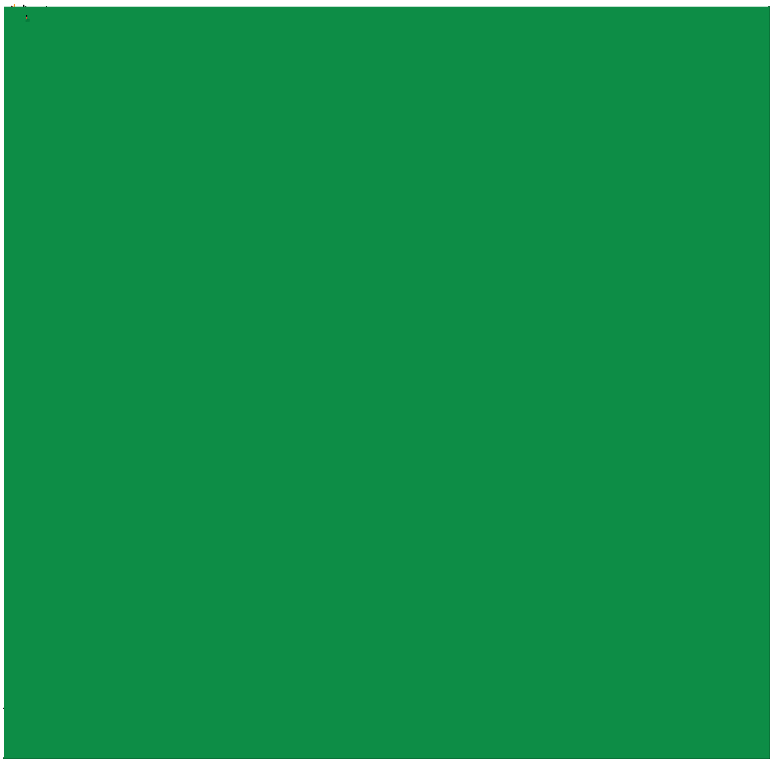

Figure 3 Error bar plot of EQ-5D score by subgroup for those receiving Aza C using intention to treat analysis

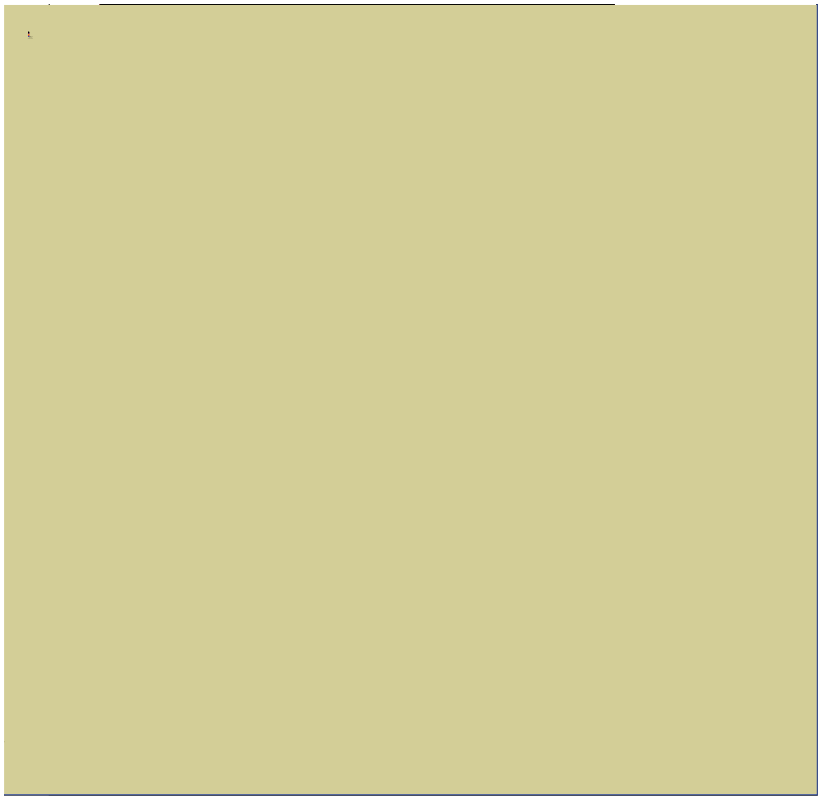

Figure 4 Error bar plot of EQ-5D score by treatment and assessment using censored analysis

Figure 5 Error bar plot of EQ-5D score by subgroup for those receiving supportive care using censored analysis

Figure 6 Error bar plot of EQ-5D score by subgroup for those receiving Aza C using censored analysis

Table 7 SF-6D utility values for different treatments using the individual level dataset used in Sekeres et al. (2004)

	Treatment									
	IC					*NIC*				
Time period	*n*	*Mean*	*SD*	*Minimum*	*Maximum*	*n*	*Mean*	*SD*	*Minimum*	*Maximum*
Baseline	21	.66	.13	.44	.86	21	.67	.08	.56	.86
2 weeks	19	.61	.10	.38	.76	19	.70	.09	.57	.86
6 weeks	18	.66	.10	.45	.86	18	.71	.15	.45	1.00
10 weeks	15	.69	.12	.55	.92	13	.72	.13	.50	1.00
14 weeks	12	.72	.16	.42	.94	8	.70	.06	.62	.80
6 months	5	.74	.18	.44	.92	4	.85	.08	.74	.92
1 year	5	.83	.10	.71	.92	4	.67	.22	.37	.86

Table 8 Mapped EQ-5D utility values for different treatments using the individual level dataset used in Sekeres et al. (2004) and the Franks et al. (2004) mapping approach[3]

	Treatment									
	IC					*NIC*				
Time period	*n*	*Mean*	*SD*	*Minimum*	*Maximum*	*n*	*Mean*	*SD*	*Minimum*	*Maximum*
Baseline	21	.59	.18	.28	.88	18	.62	.14	.39	.80
2 weeks	18	.50	.19	.09	.71	19	.66	.14	.36	.94
6 weeks	18	.57	.15	.25	.79	18	.64	.24	.11	.96
10 weeks	14	.64	.15	.30	.89	13	.68	.22	.25	.98
14 weeks	11	.66	.21	.13	.91	8	.66	.15	.43	.91
6 months	5	.73	.25	.31	.94	4	.80	.12	.70	.94
1 year	5	.80	.14	.65	.97	4	.55	.35	.05	.82

Table 9 Mapped EQ-5D utility values for different treatments using the individual level dataset used in Sekeres et al. (2004) and the Gray et al. (2006) mapping approach

	Treatment									
	IC					*NIC*				
Time period	*n*	*Mean*	*SD*	*Minimum*	*Maximum*	*n*	*Mean*	*SD*	*Minimum*	*Maximum*
Baseline	21	.62	.25	.03	.88	21	.63	.23	.09	1.00
2 weeks	19	.55	.32	.03	1.00	19	.77	.18	.36	1.00
6 weeks	18	.72	.21	.19	1.00	18	.69	.37	-.32	1.00
10 weeks	15	.69	.18	.29	1.00	13	.86	.15	.62	1.00
14 weeks	12	.71	.35	-.32	1.00	8	.76	.13	.62	1.00
6 months	5	.84	.16	.66	1.00	4	.77	.04	.73	.81
1 year	5	.87	.13	.69	1.00	4	.55	.55	-.24	1.00

[3] 'N' is reduced as the Franks et al. model requires complete SF-12 responses.

Figure 7 Error bar plot of SF-6D score by treatment using the individual level dataset used in Sekeres et al. (2004)

Figure 8 Error bar plot of EQ-5D score by treatment using the individual level dataset used in Sekeres et al. (2004) and the Franks et al. (2004) mapping approach

Figure 9 Error bar plot of EQ-5D score by treatment using the individual level dataset used in Sekeres et al. (2004) and the Gray et al. (2004) mapping approach

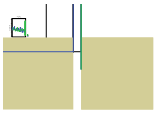

Table 10 Summary of results

	Supportive care/NIC	*Aza C*	*IC*
Phase 1			
Change in EQ-5D score from baseline to third follow up (day 182)	-0.016	0.059	
Phase 2			
Change in EQ-5D score from first to fourth assessment	0.05	0.13	
Phase 3			
Change in SF-6D score from baseline to 14 weeks	0.03		0.06
Change in EQ-5D score using Franks et al. approach from baseline to 14 weeks	0.04		0.07
Change in EQ-5D score using Gray et al. approach from baseline to 14 weeks	0.13		0.09

Note: Treatment defined as supportive care for phases 1 and 2 and non-intensive chemotherapy/best supportive care (NIC) in phase 3