International Journal of Technology Assessment in Health Care, 30:1 (2014), 90–97.

doi:10.1017/S0266462313000615

COMPARATIVE COST-EFFECTIVENESS MODELS FOR THE TREATMENT OF MULTIPLE MYELOMA

Keith Cooper, Joanna Picot, Jackie Bryant, Andrew Clegg
Southampton Health Technology Assessments Centre (SHTAC), University of Southampton

Objectives: To compare cost effectiveness models for the first-line treatment of multiple myeloma, and explore the differences between the models' structure, parameters, assumptions and results.
Methods: Three cost effectiveness models for the treatment of multiple myeloma, were compared that had been developed to inform resource allocation in the UK for the chemotherapy regimens bortezomib, melphalan and prednisolone (BMP); and melphalan, prednisolone and thalidomide (MPT) versus melphalan and prednisolone (MP). The models used alternative approaches and assumptions to estimate the overall survival and progression-free survival for each of the interventions. Through the use of sensitivity analyses, the most influential parameters and assumptions of each of the models were identified.
Results: The models developed by the manufacturers gave conflicting results, with each manufacturer favouring their drug. The differences between the model results were determined by two parameters: the hazard ratio for overall survival for MPT vs. MP and the cost of bortezomib.
Conclusions: Using models developed for assessing treatments for multiple myeloma we demonstrated that it was feasible to compare models, which then aided decision makers in making reimbursement decisions.

Keywords: Cost-utility analysis, Thalidomide, Bortezomib, Multiple myeloma

Healthcare decision makers and reimbursement agencies are continuously evaluating new healthcare technologies, by assessing clinical and cost-effectiveness data. These organizations, such as the National Institute for Health and Care Excellence (NICE) in the United Kingdom, make recommendations on the funding of new medical technologies on the basis of this evidence. Decision analytic models form the basis of the economic evaluation of these technologies as they are able to synthesize evidence on health consequences and costs from many different sources, and link intermediate outcomes from trial data to long term survival.

The NICE multiple technology appraisal (MTA) process considers evidence from an independent assessment group, together with submissions from the manufacturers of the health technology. In order for these models to be helpful in the decision-making process, it is necessary for them to have credibility and validity (1). To ensure methodological quality of the submitted models, they are assessed against the requirements of a checklist for methodological quality and generalizability to the United Kingdom (2). However, even in cases where models have adhered to these requirements for methodological quality, there are often still large differences in the results and conclusions between the models. Differences between models can be due to a combination of differences in parameter values, methods and structures and this can make between-model comparison problematic (1).

Although many guidelines exist for the development of cost-effectiveness models, and these are continually developed and refined, relatively little guidance has been developed for the assessment of model structure and assumptions, except that these should be described clearly and justified (2). In this article, we demonstrate the issues surrounding the choice of model structures and assumptions through the use of a comparative study of models developed for the evaluation of first-line treatment of multiple myeloma. These models were used to inform the guidance developed by NICE for these treatments (3). We compare the results from the independent assessment group model (the authors of this article) with those developed by the manufacturers and evaluate any differences between the models' structure, parameters, assumptions and results.

TREATMENTS FOR MULTIPLE MYELOMA

Multiple myeloma (MM) is the second most common hematological cancer in the UK, characterized by unregulated plasma cell proliferation. Myeloma is not curable, but can be treated

This work was funded by the NIHR HTA Program as project number 08/96/01. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Department of Health. We thank members of our advisory group who provided expert advice and comments. We also thank Karen Welch, Information Scientist, SHTAC, University of Southampton, for generating and running the literature searches. *Currency:* Results have been converted from UK pounds to euros for this article (with exchange rate GBP1 = EUR1.2). K. Cooper developed the research protocol, synthesized evidence, developed the economic evaluation, completed the analysis and drafted and edited the article; J. Picot developed the research protocol, managed the project, synthesized evidence, assisted in drafting the article; J. Bryant developed the research protocol, synthesized evidence, assisted in drafting the article; A.J. Clegg developed the research protocol, synthesized evidence, assisted in drafting the article.

with a combination of supportive measures and chemotherapy. The aim is to extend the duration and quality of survival by alleviating symptoms and achieving disease control while minimizing the adverse effects of the treatment. Survival of patients from diagnosis can vary from months to over a decade.

In the United Kingdom, the choice of first-line treatment depends on a combination of factors. The majority of patients are not able to withstand intensive treatment, such as high-dose chemotherapy (HDT) with autologous stem-cell transplantation (SCT), because of age or poor performance status. These patients are therefore offered single agent or combination chemotherapy which is less intensive. This study concerns the use of more recent combination therapies that incorporate drugs such as thalidomide and bortezomib (4). The three cost-effectiveness models were developed to compare the cost-effectiveness estimates of bortezomib in combination with melphalan and prednisolone (BMP) and thalidomide in combination with melphalan and prednisolone (MPT) versus melphalan and prednisolone (MP) for the first-line treatment of MM.

METHODS

The models were developed by us (the independent assessment Group, SHTAC) and the manufacturers of the two drugs under consideration (Celgene and Janssen-Cilag). The models are similar in structure, with each being a lifetime model with 6-week cycles and health states that include pre- and postprogression, and death. Health-related quality of life (HRQoL) is incorporated into the models for each of the health states and the models estimate lifetime quality-adjusted life-years (QALYs). The perspective of the analyses was that of the National Health Service (NHS) and Personal Social Services (PSS) in the United Kingdom. The models estimated the lifetime costs and benefits of treatment with discount rates of 3.5 percent, and a base price year for the costs of 2009. We describe the model developed by the authors of this study below (SHTAC model) and then outline the main differences in model structure, and assumptions used in the manufacturers' models. The results from each of the models are compared and analyzed with respect to the differences in model structure and assumptions. Results have been converted from UK Pounds to Euros for this article (with exchange rate GBP1 = EUR1.2).

DESCRIPTION OF SHTAC MODEL

The SHTAC model was used to compare the cost-effectiveness of BMP and MPT versus MP for the first-line treatment of MM (5). The model used a survival analysis approach to estimate the overall survival (OS) and progression-free survival (PFS) for each of the interventions for a patient with newly diagnosed MM. The parameter values used in the model are shown in Table 1. The model was for the duration of trial follow-up and an exponential distribution was used to extrapolate beyond the length of the trial, that is, after 36 months. Second-line treatment costs were included.

Two survival curves were constructed for OS and PFS, based on the derived probability of death and progression respectively in each model cycle. The mean survival time for OS and PFS was calculated from the survival curves for OS and PFS, using the area-under-the-curve method. The difference between the two curves provides a direct estimate of the mean time alive following disease progression until death (Figure 1). Survival was classified into three health states: *Treatment* is the time patients are treated with first-line therapy, *posttreatment* is the mean time from end of first-line treatment therapy until disease progression and *postprogression* is the mean time from disease progression until death.

The clinical parameters for the models were derived from a systematic review of the clinical effectiveness of bortezomib and thalidomide (full details given elsewhere) (5). Three studies were identified that compared MPT with MP (10–12) and one study was identified that compared BMP with MP (14). The data from the MP arms of the randomized controlled trials (RCTs), identified by the systematic review, were pooled to form baseline MP OS and PFS curves. The OS and PFS curves were derived from the studies included in the systematic review of clinical effectiveness (5).

Each health state is associated with a health related quality of life (HRQoL) utility estimate which is multiplied by the length of time spent in that state. The total QALYs over the life time of a patient is calculated by aggregating the estimated QALYs from each health state. Values for HRQoL were estimated for the treatment and posttreatment period and for those with complete response by mapping quality of life values from the European Organization for Research and Treatment of Cancer QoL questionnaire C-30 (EORTC QLQ-C30) to the EQ-5D for a cohort of MM patients receiving MP (8).

The costs in the model comprise drug treatment, consultation, and monitoring costs and costs for treating adverse events (AEs). Patients remain on drug treatment for the full treatment course unless their disease progresses or they die. Drug unit costs and doses were based on the British National Formulary 2009 (15). The duration of treatment was based on recommendations from the relevant Summary of Product Characteristics (6;7), expert clinical opinion and the published trials. The duration of treatment varies between eight cycles for MPT to nine cycles for BMP. Bortezomib is administered several times per cycle as an intravenous injection, made from a single 3.5-mg vial (7). All patients who remain alive receive second-line therapy and this is assumed to start at the mean time of disease progression for the cohort. Second-line treatment consists of either bortezomib and high dose dexamethasone (HDD); cyclophosphamide, dexamethasone, and thalidomide (CTDa); or HDD. Third-line therapy was not included as it was assumed that most patients would receive lenalidomide, irrespective of

https://doi.org/10.1017/S0266462313000615 Published online by Cambridge University Press

Table 1. Parameters and Distributions for the SHTAC Cost-Effectiveness Evaluation

Name	Mean	Range	Distribution	Reference
Cycles of treatment				
MP	8	7, 9	log normal	Clinical opinion
MPT	8	7, 9	log normal	(6)
BMP	9	8, 10	log normal	(7)
Utility values				
Treatment	0.58	0.522, 0.639	beta	(8;9)
Response	0.72	0.648, 0.792	beta	Unpublished observation
Post-progression	0.68	0.612, 0.748	beta	(8;9)
Clinical effectiveness				
Hazard rate, MP[a]				
OS baseline curve	0.028	0.020, 0.039	log normal	(10–14)
PFS baseline curve	0.067	0.060, 0.070	log normal	(10–12;14)
Hazard ratio[a]				
OS MPT vs. MP	0.62	0.50, 0.82	log normal	(10;11)
OS BMP vs. MP	0.62	0.51, 0.83	log normal	(14)
PFS MPT vs. MP	0.58	0.49, 0.77	log normal	(10;11)
PFS BMP vs. MP	0.58	0.48, 0.76	log normal	(14)
Proportion with complete response				
MP	0.026	0.017, 0.035	beta	(10–14)
MPT	0.142	0.066, 0.307	beta	(10;11)
BMP	0.217	0.121, 0.386	beta	(14)
Second-line treatment, BMP, %[b]				
MP first-line	70	60, 80	log normal	Assumption
MPT first-line	70	60, 80	log normal	Assumption
BMP first-line	15	5, 25	log normal	Assumption
Costs				
Cost of treatments, per cycle, EUR				
MP	20.26	16.20, 24.31	gamma	(15)
MPT	1,695.60	1,356, 2,034	gamma	(15)
BMP	5,303.08	4,242, 6,363	gamma	(15)
Adverse event cost, per cycle, EUR				
MP	54.76	38.33, 71.18	gamma	(16)
MPT	84.48	59.14, 109.82	gamma	(16)
BMP	87.79	61.45, 114.13	gamma	(16)
Other Costs, EUR				
Cost of bortezomib administration	184.08	128.86, 239.30	gamma	(17)
OP appointment medical oncology	145.33	101.74, 188.93	gamma	(17)
Monitoring tests at each OP visit	92.90		Fixed	(18)

[a]Cumulative hazard rates and ratios at 36 months for OS and 24 months for PFS. Calculated using duration in cycles.
[b]Shows the proportion of those receiving second-line treatment, who receive bortezomib, according to their first-line treatment. For the MP and MPT groups on second line therapy that is not BMP, 15% receive CDTa or HDD; for the BMP group 70% receive CTDa and 15% HDD.
B = bortezomib; M = melphalan; P = prednisolone; T = thalidomide; OS, overall survival; PFS, progression free survival; OP, outpatient; BMP, bortezomib in combination melphalan and prednisolone; MPT, thalidomide in combination with melphalan and prednisolone; MP, melphalan and prednisolone.

https://doi.org/10.1017/S0266462313000615 Published online by Cambridge University Press

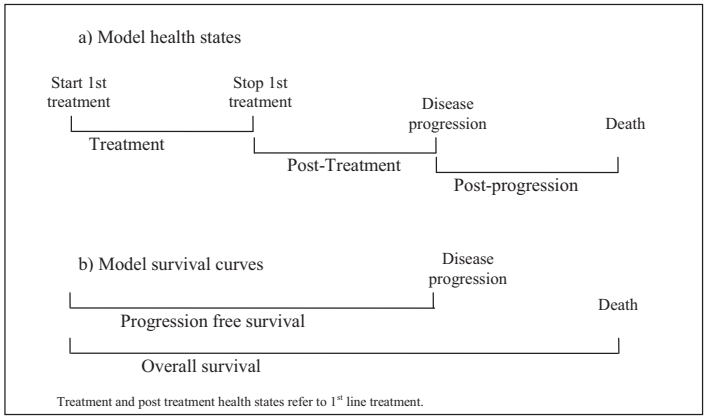

Figure 1. Schematic of the SHTAC cost-effectiveness model.

the initial treatment, as per NICE guidance (TA 171) (16). Based on clinical advice, we assumed patients attend one hospital consultation every month during their treatment period and one consultation every 3 months thereafter. The monitoring tests used for the management of MM, based on those used for the MMIX RCT, were full blood count, biochemistry, protein electrophoresis, immunoglobin, and urinary light chain excretion. For each comparator, the incidence of AEs was estimated using evidence from the RCTs included in our systematic review of clinical effectiveness (5). AEs included in the model were treatment-related serious (grade 3 and grade 4) AEs. The unit costs of treating AEs were estimated based on those used in a previous NICE technology appraisal for MM (16).

In each cycle, the total costs and QALYs are calculated by multiplying the individual costs and HRQoL by the number of people in the cohort still alive for each of the treatments. The total lifetime costs and QALYs are calculated by aggregating the costs and QALYs for all cycles. The total discounted QALY gain and cost of treatments are then calculated.

DESCRIPTION OF CELGENE MODEL

A Markov model was developed by Celgene, the manufacturer of thalidomide, to compare the difference in the progression of MM and of the costs of treatment when managed with the three different treatment options of MPT, BMP, or MP through a series of different health states

The model has four health states that are defined by the stage of disease progression or the occurrence of AEs. The four states are *pre-progression without AEs, pre-progression with AEs, post-progression* and an absorbing state of *death*. All patients start in the *pre-progression without AEs* health state and move to other states if their condition worsens or they incur an AE. Death can only occur at or after progression. The model does not include overall survival; rather it estimates the survival time before and after progression and then applies different utility values to these health states. Postprogression survival was modeled to be the same across different treatment strategies.

The model has a maximum of twelve treatment cycles for MPT and MP and nine treatment cycles for BMP. Treatment effects were calculated from a random-effects Bayesian mixed treatment comparison of data originating from three RCTs (10;11;14). The model does not include treatment costs for second or third-line therapies.

DESCRIPTION OF JANSSEN-CILAG MODEL

A survival cost-utility model was developed by Janssen-Cilag, the manufacturer of bortezomib, to compare the costs and benefits for BMP with those of MPT and MP in people with previously untreated MM who are not eligible for HDT with SCT (19). The model estimates OS and PFS curves for each of the comparators. Survival is partitioned into four different health states: before response to treatment; response but no progression; postprogression and death.

OS and PFS were estimated for MP from a meta-analysis of the MP arms of RCTs for thalidomide and bortezomib. The times to response or death were estimated from life tables constructed directly from the VISTA trial patient level data (14). For the comparator treatments, relative hazard ratios were taken from a random effects meta-analysis that used OS and PFS summary data. For estimation of the OS hazard for thalidomide, data from five RCTs were used (10;11;20–22), which included RCTs that had included thalidomide maintenance therapy.

The model includes the costs of second and third-line therapy where second-line treatment consisted of bortezomib + HDD, CTDa, or HDD. All patients receive lenalidomide plus dexamethasone as third-line treatment. HRQoL utility values are assigned to each of the states: before response to treatment, response to treatment without progression, and postprogression, based on a study evaluating chemotherapy followed by SCT in people with MM (23).

The duration of treatment with MP is seven cycles as per the VISTA trial (14). For bortezomib, 31.5 vials were used per patient based on usage in the VISTA trial, which is lower than the full treatment course of 52 vials. For thalidomide, the model

https://doi.org/10.1017/S0266462313000615 Published online by Cambridge University Press

Table 2. SHTAC and the Manufacturers' Baseline Cost-Effectiveness Results versus MP

	Analysis	MP	MPT	BMP
Total cost, EUR	SHTAC	25,727	39,118	68,602
	Janssen-Cilag	65,321	71,186	80,011
	Celgene	1,638	25,360	51,139
Total QALY	SHTAC	2.42	3.64	3.62
	Janssen-Cilag	2.86	3.41	4.03
	Celgene	2.43	3.28	3.35
Incremental cost vs MP, EUR	SHTAC	-	13,391	42,875
	Janssen-Cilag	-	5,866	14,690
	Celgene	-	23,722	49,501
Incremental QALY vs MP	SHTAC	-	1.22	1.2
	Janssen-Cilag	-	0.55	1.17
	Celgene	-	0.85	0.92
ICER vs MP, EUR per QALY	SHTAC	-	10,962	35,784
	Janssen-Cilag	-	10,694	12,598
	Celgene	-	28,057	54,029
ICER, BMP vs MPT, EUR per QALY	SHTAC	-	-	Dominated
	Janssen-Cilag	-	-	14,288
	Celgene	-	-	364,614

B, bortezomib; T, thalidomide; MP, melphalan and prednisolone; ICER, incremental cost effectiveness ratio; EUR, euro; QALY, quality-adjusted life-year; BMP, bortezomib in combination melphalan and prednisolone; MPT, thalidomide in combination with melphalan and prednisolone; MP, melphalan and prednisolone; SHTAC, Southampton Health Technology Assessments Centre.

used an average duration of treatment of 315 days based on the duration reported in the MPT RCTs.

RESULTS

Comparison of Economic Evaluation Results

The results for the manufacturers' and SHTAC's economic analyses are shown in Table 2. The different assumptions and methodology described above result in a range of estimates for the cost and benefits of the treatment options.

The incremental cost-effectiveness ratio (ICER) for MPT versus MP varies between EUR10,694 (Janssen-Cilag) and EUR28,057 (Celgene) per QALY gained. The ICER for BMP versus MP varies between EUR12,598 (Janssen-Cilag) and EUR54,029 (Celgene) per QALY gained.

The results of the analyses comparing BMP and MPT vary considerably. For BMP versus MPT, the ICER was estimated as EUR14,288 (Janssen-Cilag), and EUR364,614 (Celgene) per QALY gained. For the SHTAC economic analysis, MPT dominated BMP, that is, MPT is cheaper and more effective, for the base case analysis. Thus the conclusions differ between the analyses, with Janssen-Cilag concluding that BMP is cost-effective compared with MPT, whereas the other two analyses disagree.

The costs vary substantially between the analyses, for example the cost of MP varies between EUR1,638 for the Celgene submission and EUR65,321 for the Janssen-Cilag submission. The costs from the Celgene analysis were lower as they had not included any subsequent treatment costs, whereas the SHTAC analysis included costs for second-line treatment and the Janssen-Cilag included costs for second- and third-line treatment.

The incremental costs for MPT versus MP vary between EUR5,866 (Janssen-Cilag) and EUR23,722 (Celgene). The Celgene submission uses higher dosages of thalidomide (238 mg/day) for longer periods (eleven cycles) than the other two analyses. The incremental costs for BMP versus MP vary between EUR14,690 (Janssen-Cilag) and EUR49,501 (Celgene). These differences are largely due to the assumptions around the number of vials of bortezomib used, with Janssen-Cilag assuming a mean of 31.5 vials used per person, whereas the mean number of vials used is over forty in the SHTAC and Celgene economic evaluations.

The total QALY estimates between the studies are reasonably similar with estimates for each treatment: MP (range, 2.42–2.86), MPT (range, 3.28–3.64), and BMP (range, 3.35–4.03). The incremental QALY estimates for MPT versus MP vary widely and these differences are due to the estimates chosen for the hazard ratio for OS compared with MP, range from 0.55 (Janssen-Cilag) to 1.22 (SHTAC).

Sensitivity Analyses of the SHTAC Model

One-way deterministic sensitivity analyses were previously performed on all the parameters in the SHTAC model, described in detail elsewhere (5), and the model results were found to be most sensitive to the hazard ratio for OS, cost and dosage of the treatment, and the overall baseline survival curve used for MP. Sensitivity analyses were then performed on the SHTAC model to assess the effect of the different assumptions used between the models for the estimates of hazard ratio for OS for MPT and cost for bortezomib.

The estimate for the effectiveness of OS for MPT versus MP varies according to the sources of data chosen. The SHTAC model based its estimate on a systematic review (5), (hazard ratio 0.62). This review excluded trials in which participants had received maintenance therapy with thalidomide. In contrast, the Janssen-Cilag analysis used an OS estimate based on a meta-analysis that included trials in which patients had maintenance therapy. The inclusion of studies with maintenance therapy resulted in lower improvement in OS for MPT versus MP than the meta-analysis without studies of maintenance therapy.

The cost of BMP varies substantially between the analyses. The SHTAC model based its estimate on the number of cycles and doses specified by the VISTA trial and the Summary of Product Characteristics (7;14). In the model, this amounted to approximately forty-eight vials of treatment. In contrast, the Janssen-Cilag analysis uses a lower cost, due to a lower number

https://doi.org/10.1017/S0266462313000615 Published online by Cambridge University Press

Table 3. Deterministic Sensitivity Analyses from the SHTAC Model

SHTAC model results, EUR/QALY	MPT vs. MP	BMP vs. MP	BMP vs. MPT
Base case	10,962	35,784	MPT dominates BMP
a) Hazard ratio OS MPT = 0.8	24,196	35,784	44,294
b) Reduced cost for BMP	10,962	24,084	MPT dominates BMP
a) and b) Hazard ratio OS MPT = 0.8; and reduced cost for BMP	24,196	24,084	24,001

SHTAC, Southampton Health Technology Assessments Centre; BMP, bortezomib in combination melphalan and prednisolone; MPT, thalidomide in combination with melphalan and prednisolone; MP, melphalan and prednisolone; OS, overall survival.

of treatment doses, based on the actual number of treatment vials used in the VISTA trial, that is, 31.5 vials. The reduced number of vials may be due to fewer treatment cycles due to early discontinuation of treatment.

In the sensitivity analyses the hazard ratio for OS for MPT versus MP was varied from 0.62 (base case) to 0.8. The cost of bortezomib was varied according to the proportion of early discontinuation from treatment. In each cycle, the proportion who discontinued treatment was varied between 0 (baseline) and 10 percent. The results of the sensitivity analyses are shown in Table 3.

The sensitivity analyses show that these parameters have a large effect on the model results, when comparing BMP to MPT, while the comparison between MPT vs. MP and BMP versus MP are more robust. In the base case analysis, MPT dominates BMP, as MPT is both cheaper and more effective than BMP. With the alternative hazard ratio for overall survival, the ICER is EUR44,294 per QALY gained for BMP versus MPT. For both the alternative hazard ratio for OS and a reduced cost for BMP, the ICER is EUR24,001 per QALY gained. These results are consistent with the results estimated by each of the models with the assumptions they used.

This uncertainty within the three models was explored further by conducting a probabilistic sensitivity analysis (PSA). All parameters were sampled probabilistically using the ranges and values shown in Table 1. However, in this case the hazard ratio OS for MPT versus MP was varied between the two estimates (0.62–0.8) using a beta distribution, and the treatment discontinuation rate for BMP was varied between the two estimates (0–10 percent) using a uniform distribution to represent the uncertainty between the parameter inputs in the three models. One thousand simulations were run. The cost-acceptability curve is shown in Appendix 2 and indicates that at lower willingness to pay thresholds of between EUR24,000 and EUR60,000 per QALY gained (GBP20,000 and GBP50,000 per QALY gained) MPT has the highest probability of being cost-effective. For a willingness to pay threshold higher than EUR60,000 per QALY gained, BMP is the treatment with the highest probability of being cost-effective.

DISCUSSION

In the United Kingdom, NICE has provided guidance for the use of bortezomib and thalidomide for first-line treatment of MM, based upon the clinical and cost-effectiveness evidence (3). It recommended thalidomide, in combination with an alkylating agent and a corticosteroid, as a cost-effective option for the first-line treatment of MM, with bortezomib recommended for those unable to tolerate or who have contra-indications to thalidomide. NICE assessed evidence from the manufacturers of thalidomide and bortezomib, from our clinical and economic evaluation, and the opinion of clinical experts and service users. For the estimation of the clinical effectiveness of MPT, the NICE appraisal committee decided that it was appropriate to exclude trials in which participants received maintenance therapy with thalidomide. For the estimation of the costs of bortezomib, the NICE appraisal committee accepted the number of bortezomib vials during treatment would be 31.5 (as proposed by Janssen-Cilag).

The approach taken in this article to compare cost-effectiveness models has shown that it is possible to compare and evaluate cost-effectiveness models, by identifying the most influential differences with respect to model structure, parameter and assumptions, and making a judgment on these differences. Although this approach is necessary within a decision-making context for national regulatory bodies, such as NICE, it is not common in the medical literature. Turner et al. (1) investigated the feasibility of between-model comparison by comparing four UK models developed for coronary heart disease. They concluded that, while checking between model consistency requires a potentially large investment in terms of researcher time and effort, there were situations where it would be useful for decision makers, for example where there were large differences in model results, or where results from different models cross decision-maker's thresholds. In this case, there may be considerable uncertainty to the implications of results to decision making. They noted that often detailed information on the models was restricted by word count limitations and recommended that modeling articles should include detailed Web appendices to aid replication and checks of between-model consistency.

https://doi.org/10.1017/S0266462313000615 Published online by Cambridge University Press

In the current article, comparison was aided as each of the models involved had been constructed to conform to NICE guidelines for technology appraisals (24), and full details of the economic models were available and the authors had access to electronic versions of the models. Although there were many differences between the model results, it was possible to identify those parameters that primarily caused the differences and examine them for validity. Through analyzing these differences, we concluded that each modeling approach and structure was appropriate and that many of the differences between the models in terms of parameter inputs and assumptions have negligible effect on the model results.

Haji Ali Afzali and Karnon (25) propose the concept of reference models for specific disease areas, which could be made available to sponsors submitting health technologies for assessment by reimbursement bodies. These resulting models would be more likely to represent a comprehensive, unbiased representation of the disease. They argue that there is a diversity of model structures within disease areas which increases the complexity of comparing evaluations of alternative technologies for the same conditions. Furthermore, they consider that the consequences of inconsistencies in the choice of model structure within a specific disease for submissions to a national regulatory body can lead to inconsistent reimbursement decisions because changes in model structure and analysis can produce substantially different results. The comparison of existing models is a useful method to reach consensus between modelers. For example, the Mount Hood Challenge (26) was a forum for computer modelers of diabetes to discuss and compare models and identify key areas of future development to advance the field. By performing systematic comparisons and validation exercises enabled the identification of key differences among the models, as well as their possible causes and directions for improvement in the future.

CONCLUSIONS

A comparison of alternative cost-effectiveness models is not straightforward, and often differences in model structure, parameters, and assumptions are hard to identify and lead to large differences in results and conclusions. By comparing models developed for assessing treatments for multiple myeloma, we demonstrated that it was feasible to compare models, which then aided decision makers in making reimbursement decisions.

CONTACT INFORMATION

Keith Cooper, PhD (k.cooper@soton.ac.uk), Senior Research Fellow, Southampton Health Technology Assessments Centre (SHTAC), University of Southampton, UK

Joanna Picot, PhD, Senior Research Fellow, SHTAC, University of Southampton, UK

Jackie Bryant, MSc, Principal Research Fellow, SHTAC, University of Southampton, UK

Andrew Clegg, PhD, Professor of Health Services Research, and Director of SHTAC, University of Southampton, UK

CONFLICTS OF INTEREST

All authors report they have received a grant to their institution from National Institute for Health Research to undertake research.

APPENDIX 1

Cost-acceptability curve from the probabilistic sensitivity analysis from the SHTAC model (Appendix Figure 1) available online at http://dx.doi.org/10.1017/S0266462313000615.

REFERENCES

1. Turner D, Raftery J, Cooper K, et al. The CHD challenge: Comparing four cost-effectiveness models. *Value Health.* 2011;14:53-60.
2. Philips Z, Ginnelly L, Sculpher M, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess.* 2004;8:iii-xi, 1.
3. National Institute for Health and Clinical Excellence. *Bortezomib and thalidomide for the first-line treatment of multiple myeloma*. London: NICE; 2011.
4. Messori A, Maratea D, Nozzoli C, Bosi A. The role of bortezomib, thalidomide and lenalidomide in the management of multiple myeloma: An overview of clinical and economic information. *Pharmacoeconomics.* 2011;29:269-285.
5. Picot J, Cooper K, Bryant J, Clegg A. The clinical and cost-effectiveness of bortezomib and thalidomide for the first-line treatment of multiple myeloma: A systematic review and economic evaluation. *Health Technol Assess.* 2011;15:1-204.
6. Summary of Product Characteristics for Thalidomide Celgene. Revision 2 - Published 16/04/09. 2009. European Public Assessment Report. 22-5-2009.
7. Summary of Product Characteristics for Velcade. Revision 17 - Published 20/04/09. 2009. European Public Assessment Report. 22-5-2009.
8. Gulbrandsen N, Hjermstad, MJ, Wisloff, F, Nordic Myeloma Study Group. Interpretation of quality of life scores in multiple myeloma by comparison with a reference population and assessment of the clinical importance of score differences. *Eur J Haematol.* 2004;72:172-180.
9. McKenzie L, van der Pol M. Mapping the EORTC QLQ C-30 onto the EQ-5D Instrument: The potential to estimate QALYs without generic preference data. *Value Health.* 2009;12:167-171.
10. Facon T, Mary JY, Hulin C, et al. Melphalan and prednisone plus thalidomide versus melphalan and prednisone alone or reduced-intensity autologous stem cell transplantation in elderly patients with multiple myeloma (IFM 99-06): A randomised trial. *Lancet.* 2007;370:1209-1218.
11. Hulin C, Facon T, Rodon P, et al. Efficacy of melphalan and prednisone plus thalidomide in patients older than 75 years with newly diagnosed multiple myeloma: IFM 01/01 Trial. *J Clin Oncol.* 2009;27:3664-3670.
12. Palumbo A, Bringhen S, Caravita T, et al. Oral melphalan and prednisone chemotherapy plus thalidomide compared with melphalan and prednisone alone in elderly patients with multiple myeloma: Randomised controlled trial. *Lancet.* 2006;367:825-831.
13. MRC Myeloma IX Trial Outcomes: Results from the non transplant (Non-intensive) group. Academic in Confidence Data. 2009.
14. San Miguel JF, Schlag R, Khuageva NK, et al. Bortezomib plus melphalan and prednisone for initial treatment of multiple myeloma. *N Engl J Med.* 2008;359:906-917.

15. BMA, Royal Pharmaceutical Society of Great Britain. *British National Formulary 57*. London: BMJ Publishing Group; 2009.
16. National Institute for Health and Clinical Excellence. *Lenalidomide for the treatment of multiple myeloma in people who have received at least one prior therapy (TA 171)*. London: NICE; 2009.
17. Department of Health. *NHS reference costs*. 2008. UK.
18. Contracting Unit Southampton University Hospital Trust. Unit costs. 2009. (unpublished database).
19. Janssen-Cilag. *Bortezomib and thalidomide for the first-line treatment of multiple myeloma: Bortezomib (Velcade)*. Executive summary. 1-4. 2009. London: National Institute for Health and Clinical Excellence.
20. Gulbrandsen N, Waage A, Gimsing P, et al. A randomised placebo controlled study with melphalan/prednisone vs melphalan/prednisone/thalidomide: Quality of life and toxicity. *Haematologica.* 13th Congress of the European Hematology Association, Copenhagen, Denmark. 2008;93.
21. Palumbo A, Bringhen S, Liberati AM, et al. Oral melphalan, prednisone, and thalidomide in elderly patients with multiple myeloma: Updated results of a randomized controlled trial. *Blood.* 2008;112:3107-3114.
22. Wijermans PW, Zweegman S, van Marwijk KM, et al. MP versus MPT in elderly myeloma patients: The final outcome of the HOVON-49 study (Abstract A116). *Clin Lymphoma Myeloma.* 2009;9:S18.
23. van Agthoven M, Segeren CM, Buijt I, et al. A cost-utility analysis comparing intensive chemotherapy alone to intensive chemotherapy followed by myeloablative chemotherapy with autologous stem-cell rescue in newly diagnosed patients with stage II/III multiple myeloma: A prospective randomised phase III study. *Eur J Cancer.* 2004;40:1159-1169.
24. National Institute for Health and Clinical Excellence. *Guide to the methods of technology appraisal*. London: NICE; 2008.
25. Haji Ali Afzali AH, Karnon J. Addressing the challenge for well informed and consistent reimbursement decisions: The case for reference models. *Pharmacoeconomics.* 2011;29:823-825.
26. Computer modeling of diabetes and its complications: A report on the Fourth Mount Hood Challenge Meeting. *Diabetes Care.* 2007;30:1638-1646.

https://doi.org/10.1017/S0266462313000615 Published online by Cambridge University Press