Target journal: Pharmacoeconomics

Article type: Original article or systematic review

Title: Economic evaluations of pharmaceuticals without randonmised controlled trials: a systematic review and taxonomy

Anthony J Hatswell MSc\textsuperscript{1,2}, Nick Freemantle PhD\textsuperscript{3}, Gianluca Baio PhD\textsuperscript{1}

1 Department of Statistical Science, University College London, Gower Street, WC1E 6BT
2 BresMed, 84 Queen Street, Sheffield, S1 2DW
3 Department of Primary Care and Population Health, University College London, Gower Street, WC1E 6BT

Key words: historical control; non-randomised; single group study, health technology appraisal, cost-effectiveness
Abstract (247 / 250)

Background
Pharmaceuticals are usually granted a marketing authorisation on the basis of randomised controlled trials (RCTs). Occasionally the efficacy of a treatment is assessed without a randomised comparator group (either active or placebo).

Objective
To identify and develop a taxonomic account of economic modelling approaches for pharmaceuticals licensed without randomised control trial data.

Methods
We searched PubMed, the websites of UK health technology assessment (HTA) bodies and the ISPOR Scientific Presentations Database for assessments of the 74 indications granted a marketing authorisation by the FDA or EMA from Jan-1999 to May-2014 without RCT data. The outcome of interest was the approach to modelling efficacy data.

Results
Fifty-one unique models were identified based on 29 peer-reviewed articles, 30 health technology appraisals (HTAs), and 15 ISPOR abstracts concerning 30 indications (44 indications had not been modelled). Also notable was the low rate of submission to HTA agencies (28/98). The majority of models (43/51) were based on ‘historical controls’ – comparisons to previous meta-analysis or pooling of trials (5), individual trials (19), registries/case series (14), or expert opinion (7). Other approaches used the patient as their own control, performed threshold analysis, assumed time on treatment was added to overall survival, or performed cost-minimization.

Conclusions
There is considerable variation in the quality and approach of models constructed for drugs granted a marketing authorisation without a RCT. The most common approach is of a naïve comparisons to historical data (using other trials / registry data as a control group), with considerable scope for bias.
Key points for decision makers

- When pharmaceuticals are licensed without comparative data, economic models are generally constructed using a historical control.
- Even within evaluations using the same method, the quality and appropriateness of an approach varies.
- The appropriateness of historical comparison for modelling using uncontrolled study data requires further methodological evaluation.
Article: Word count: 2261

1. Introduction

Treatments are usually granted a marketing authorisation on the basis of randomised controlled trials, conducted against either a placebo or an active control[1]. This provides a good basis for regulators to make decisions regarding the efficacy of interventions compared to the current standard of care[2]. This evidence may then be used to estimate the difference between the new treatment and the standard of care. Indirect treatment comparisons using a common comparator sometimes enable the comparison of the efficacy of treatments in different studies[3,4].

Less commonly, treatments can be granted a marketing authorisation without a study containing a control arm. In a few cases it may be ‘obvious’ that the treatment is efficacious, for example if all patients died before an intervention was available, but all live afterwards[5], or patients achieve a marked improvement in an objective measure, for example blood count[2]. Whilst these treatments may receive a license for their use, estimates of their comparative efficacy compared to the currently used treatment are still needed to inform decisions on reimbursement in many healthcare systems. This may be to show the clinical improvement for patients, or alternatively for cost-effectiveness analysis, used in many countries for resource allocation decisions[3].

Where cost-effectiveness analysis is used as a criterion in healthcare systems, treatments are required to generate more health (usually defined in terms of quality adjusted life years) than the treatments that would be displaced (represented by a ‘shadow budget’), i.e. that the money spent on the new intervention would not be better spent elsewhere. How such estimates should be constructed without controlled trials is however unclear – whilst there exists extensive guidance on constructing economic models based on RCT results, there is no guidance on the most appropriate method of modelling uncontrolled study data from health technology agencies or professional bodies (Table 1).

The objective of this study was therefore to identify models constructed for treatments granted marketing authorisation without controlled clinical trial evidence, and the approach taken to estimating relative efficacy between the treatment(s). Of relevance to the study were both published economic models and health technology appraisals.

2. Methods

Hatswell et al. [6] identified 74 indications granted a marketing authorisation by either the FDA or the EMA from 1999 to May 2014, on the basis of uncontrolled clinical trial results. We conducted a systematic review for economic evaluations published for each of the treatments listed using PubMed (search terms given in Figure 1). Searches were also conducted for health technology appraisals conducted by the National Institute for Health and Care Excellence (NICE), the Scottish Medicines Consortium (SMC), and All Wales Medicines Strategy Group (AWMSG), as well as the grey literature of the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Scientific Presentations Database.

After identification, results (papers, health technology appraisal submissions and scientific presentations) were filtered for models constructed in the relevant indications with uncontrolled
study data (some treatments had multiple indications, which was what the publication investigated, or subsequently had RCTs conducted). Results were then de-duplicated based on the model descriptions and study authors, to account for the same model being used for different purposes (for example a model used in a NICE submission, then published with Spanish costs, all whilst using the same approach to modelling efficacy). Where it was not clear whether a model was reported on multiple occasions, or was a similar approach, this was discussed by the reviewers and a decision reached by consensus.

Following identification of the economic models, the approaches used to estimate efficacy against the relevant comparator were categorised for each model. If a model included multiple approaches to modelling efficacy data, these were classed as separate approaches. The resulting models were then placed in to a taxonomic framework and analysed for commonality in approach.

3. Results

Figure 2 shows the PRISMA diagram for economic evaluations retrieved through PubMed [7]. The initial 74 searches yielded 1202 hits, which were reduced to 56 full articles after abstract and title review. Twenty nine papers were included in the final analysis. The main reasons for exclusion on full paper review were models being based on RCT data (n=9), considering a different indication (including a different stage of the same disease, n=7) and papers that did not contain an economic model (e.g. burden of illness studies, n=6).

In addition to published papers searches of health technology body websites led to 19 NICE appraisals being identified (9 included), 52 SMC appraisals identified (16 included), and 27 AWMSG appraisals identified (5 included). Overall there was a notable level of non-submission to HTA agencies, in particular to the SMC (13/52 non-submissions) and AWMSG (13/27 non-submissions). Appraisals also often occurred after RCT based results had become available (NICE 8/19, SMC 9/52, AWMSG 3/27) leading to exclusion from this study. Full results of the review are shown in Table 1.

Searching the ISPOR Scientific Presentations Database led to 1780 abstract hits, with 43 records selected for further review and 29 full records included. The most common reason for exclusion was insufficient information reported regarding the model or approach used (n=14).

In total, 74 relevant documents were identified (including publications, health technology appraisals and scientific presentations), which described 91 distinct modelling approaches. After consolidation of approaches reported multiple times (for example one model being used for NICE and SMC submissions, presented at ISPOR and then published in an indexed journal), 51 unique approaches were identified. Of these 51 models, the overwhelming majority (n=43, 84%) were based on historical controls. Other approaches identified included using patients as their own control either through statistical analysis or comparisons with baseline values (n=3, 6%), cost minimisation analyses (n=3, 6%), threshold analyses (n=1, 2%), or assuming in oncology that time in progression free survival was added to overall survival, with treatments then given in sequence (termed the ‘cumulative method’; n=1, 2%) (Figure 3).

All the 43 historical controls identified compared the results of the uncontrolled study of the new treatment to a separate set of data. In 17 cases (40%) the new treatment was compared to one arm
from another clinical study and in 5 cases (12%) pooled or meta-analysed data from a series of studies. A further 14 models (33%) used comparisons to registry or case series data, with 7 models (16%) comparing the results of the uncontrolled study to expert opinion. Trial and registry data appeared to be used interchangeably in evaluations, with only seven studies (16%) attempting to account for differences in patient characteristics or patient selection.

4. Discussion

The results of this review show that 51 unique models have been published for 30 different indications granted a marketing authorisation without a comparative trial. Consequently of the 74 indications approved without a comparative trial, [6] 44 indications have therefore not been modelled and estimates of cost-effectiveness are available. It is not known what the rate of economic evaluation of new indications is, although we suspect it will be higher than the 40% rate seen in this study.

The use of historical control was by far the most common approach (43/51), which was most frequently taken from another trial or trials (22 of the 43). However, even within this method there was substantial variation – some studies compared the results of uncontrolled trials to results taken from multiple trials (for example Dinnes et al. who pooled the results of 8 other clinical trials to compare against), whereas the majority of models compared against single arms from other studies.

The assumption inherent in naïve comparisons to historical controls (first proposed by Pocock[8]) is that patients are similar, or “exchangeable”, between studies. If this is not the case, and patient systemically differ, then this procedure will introduce bias in the results. Several approaches to matching patients and baseline characteristics between studies are available in the literature, including methods based on propensity scores,[9] and match adjusted indirect comparisons [10]. Despite the availability of these approaches, only 2 models attempted to control for any differences between trials, with one notable outlier being the work by Annemans et al. [11], who constructed a historical control by reviewing patient records at the centres that participated in the clinical trial, in the time period before the clinical trial was open for enrolment [11].

The lack of adjustment of outcomes to reflect potentially more favourable patient cohorts may represent a substantial bias in the literature in favour of the new treatments. A study by Sacks, Chalmers & Smith of 50 RCTs and 56 historically controlled trials of the same interventions, the randomised control arm performed better than the historical control arm. In the studies cited therefore 79% of historically controlled trials stated the intervention was effective, compared to only 20% of RCTs[12]. Diehl & Perry investigated the same question looking at overall survival or relapse free survival in oncology, finding 43 examples in the literature of well-matched historical cohorts and RCT control groups. However when comparing the outcomes of the two groups, 18 of the 43 studies had a greater than 10% difference in effect size between the control groups – the randomised group performing better on 17/18 occasions [13]. This latter finding is particularly concerning given 32 of the 43 historically controlled models were in oncology, though other examples historical controls have proved a poor match for RCT control arms that would be expected to have shown similar results. [14–16]
Outside of historical controls, cost minimisation (though frowned upon in the literature[17]) was used in 3 models – whilst it may appear superficially attractive to assume treatments have equal efficacy to similar ones, it is unlikely that they exhibit exactly the same efficacy, with zero uncertainty. A further 3 models compared patient outcomes on treatment to a patient’s baseline result. This is also a potentially biased approach, due to issues such as regression to the mean[18]. One additional approach, comparing all patients to non-responders, allows the estimation of an effect size; however it will be overly favourable towards the intervention, as non-responders will include an inherently sicker population [19]. The final approach noted was that of Tappenden et al., who pragmatically performed threshold analysis of the relative risk needed for the drug to be considered cost-effective. Although this does not necessarily give an estimate of effect size, it allows a decision maker to make a more informed decision after reviewing the clinical evidence[20], as such we would recommend the use of similar threshold analyses where appropriate.

That there is a number of differing approaches to modelling, with a lack of standard approach to handling issues such as patient selection, is likely a reflection of the relative rarity of evaluations with this type of data (we identified only 51 models, compared to the vast literature of health economic evaluations published[21]). Despite the lack of standard approaches and guidelines however, some studies appear to be well conducted, with attempts to select an approach based on reasonable assumptions, and control for any patient selection (for example Woods et al.[22]).

5. Conclusion

The majority of treatments granted a marketing authorisation without controlled study results have not been subject to economic evaluation in a published form, with also a high level of non-submission to UK health technology agencies for such products. The evaluations that have been performed were generally based on naive comparisons to historical controls from individual arms of clinical trials, or registry/case series data.

Further research and guidance is required on the appropriateness of historical controls in economic evaluation, and on the most relevant methods to use when modelling uncontrolled data with the aim of estimating comparative effectiveness. Ultimately formal guidance and standardisation may reduce the level of bias in economic evaluations, and lead to an improvement in the average quality of published models. Standardisation would also provide a basis for comparison between studies, such that interventions can be more readily compared with other approaches to evaluation, where methods are comparable [23].
References


47. Merck Pharmaceuticals. Cetuximab (Erbitux) 100mg solution for infusion (2mg/mL) - Submission to the National Institute for Health and Clinical Excellence. National Institute for Health and Clinical Excellence; 2005.


52. AWMSG. Final Appraisal Report - Trabectedin (Yondelis) for advanced soft tissue sarcoma. All Wales Medicines Strategy Group; 2008.

53. Scottish Medicines Consortium. Trabectedin, 0.25 and 1mg powder for concentrate for solution for infusion (Yondelis). Scottish Medicines Consortium; 2010.

54. Scottish Medicines Consortium. 2nd Re-Submission In Confidence - Trabectedin, 0.25 and 1mg powder for concentrate for solution for infusion (Yondelis). Scottish Medicines Consortium; 2011.


64. AWMSG. Final Appraisal Report - Nelarabine for treatment of T-cell acute lymphoblastic leukaemia and T-cell lymphoblastic lymphoma. All Wales Medicines Strategy Group; 2009.


77. AWMSG. Final Appraisal report - dasatinib (Sprycel®) Chronic, accelerated or blast phase CML. All Wales Medicines Strategy Group; 2007.


85. Scottish Medicines Consortium. Tocofersolan, 50mg/mL (corresponding to 74.5 IU tocopherol) oral solution (Vedrop®) SMC No. (696/11). Scottish Medicines Consortium; 2012.


89. Hoyle M, Crathorne L, Moxham T, Garside R, Hyde C. Ofatumumab (Arzerra®) for the treatment of chronic lymphocytic leukaemia in patients who are refractory to fludarabine and alemtuzumab: a critique of the submission from GSK. University of Exeter; 2010.


47. Merck Pharmaceuticals. Cetuximab (Erbitux) 100mg solution for infusion (2mg/mL) - Submission to the National Institute for Health and Clinical Excellence. National Institute for Health and Clinical Excellence; 2005.


52. AWMSG. Final Appraisal Report - Trabectedin (Yondelis) for advanced soft tissue sarcoma. All Wales Medicines Strategy Group; 2008.

53. Scottish Medicines Consortium. Trabectedin, 0.25 and 1mg powder for concentrate for solution for infusion (Yondelis). Scottish Medicines Consortium; 2010.
54. Scottish Medicines Consortium. 2nd Re-Submission In Confidence - Trabectedin, 0.25 and 1mg powder for concentrate for solution for infusion (Yondelis). Scottish Medicines Consortium; 2011.


64. AWMSG. Final Appraisal Report - Nelarabine for the treatment of T-cell acute lymphoblastic leukaemia and T-cell lymphoblastic lymphoma. All Wales Medicines Strategy Group; 2009.


77. AWMSG. Final Appraisal report - dasatinib (Sprycel®) Chronic, accelerated or blast phase CML. All Wales Medicines Strategy Group; 2007.


85. Scottish Medicines Consortium. Tocofersolan, 50mg/mL (corresponding to 74.5 IU tocopherol) oral solution (Vedrop*) SMC No. (696/11). Scottish Medicines Consortium; 2012.


89. Hoyle M, Crathorne L, Moxham T, Garside R, Hyde C. Ofatumumab (Arzerra®) for the treatment of chronic lymphocytic leukaemia in patients who are refractory to fludarabine and alemtuzumab: a critique of the submission from GSK. University of Exeter; 2010.


### Table 1: Tabulated PRISMA diagram of the number and source of economic evaluations identified as being based on uncontrolled clinical study data

<table>
<thead>
<tr>
<th></th>
<th>NICE</th>
<th>SMC</th>
<th>AWMSG</th>
<th>ISPOR</th>
<th>PubMed</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of hits</td>
<td>19</td>
<td>52</td>
<td>27</td>
<td>1793</td>
<td>1235</td>
<td>3182</td>
</tr>
<tr>
<td>Non-submissions</td>
<td>2</td>
<td>13</td>
<td>13</td>
<td>-</td>
<td>-</td>
<td>28</td>
</tr>
<tr>
<td>For review</td>
<td>17</td>
<td>39</td>
<td>14</td>
<td>43</td>
<td>56</td>
<td>142</td>
</tr>
<tr>
<td>Excluded</td>
<td>8</td>
<td>23</td>
<td>9</td>
<td>28</td>
<td>27</td>
<td>95</td>
</tr>
<tr>
<td>Different indication</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>Clinical paper or commentary</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Not an economic model</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>Model based on RCT data</td>
<td>8</td>
<td>9</td>
<td>3</td>
<td>5</td>
<td>9</td>
<td>33</td>
</tr>
<tr>
<td>Insufficient information</td>
<td>0</td>
<td>9</td>
<td>5</td>
<td>14</td>
<td>1</td>
<td>29</td>
</tr>
<tr>
<td>Included</td>
<td>9</td>
<td>16</td>
<td>5</td>
<td>15</td>
<td>29</td>
<td>74</td>
</tr>
<tr>
<td>Drug</td>
<td>Disease area</td>
<td>Approach taken</td>
<td>Historical control?</td>
<td>Source of historical data</td>
<td>Data adjusted?</td>
<td>Year</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>---------------------</td>
<td>---------------------------</td>
<td>----------------</td>
<td>----------</td>
</tr>
<tr>
<td>Anagrelide</td>
<td>Essential thrombocytopenia</td>
<td>Decision tree comparing outcomes with treatment to patient baseline</td>
<td>No</td>
<td>-</td>
<td>-</td>
<td>1999</td>
</tr>
<tr>
<td>Anagrelide</td>
<td>Essential thrombocytopenia</td>
<td>Markov model using other expert opinion on the efficacy of treatments</td>
<td>Yes</td>
<td>Expert opinion</td>
<td>-</td>
<td>2002</td>
</tr>
<tr>
<td>Anagrelide</td>
<td>Essential thrombocytopenia</td>
<td>Markov model using historical control data from a trial for the main comparator and no treatment</td>
<td>Yes</td>
<td>Trial</td>
<td>-</td>
<td>2005</td>
</tr>
<tr>
<td>Argatroban</td>
<td>Heparin-induced thrombocytopenia</td>
<td>Trial data used for the drug, with historical data from the same hospitals involved in the trial used for the comparator</td>
<td>Yes</td>
<td>Case series</td>
<td>-</td>
<td>2006</td>
</tr>
<tr>
<td>Argatroban</td>
<td>Heparin-induced thrombocytopenia</td>
<td>Trial data used for the drug, with case series data used for the control</td>
<td>Yes</td>
<td>Case series</td>
<td>-</td>
<td>2007</td>
</tr>
<tr>
<td>Argatroban</td>
<td>Heparin-induced thrombocytopenia</td>
<td>Decision tree model using historical control data from hospitals for both treatment and control</td>
<td>Yes</td>
<td>Case series</td>
<td>-</td>
<td>2012</td>
</tr>
<tr>
<td>Argatroban</td>
<td>Heparin-induced thrombocytopenia</td>
<td>Decision tree model using an assumption of equal efficacy (cost minimisation)</td>
<td>-</td>
<td>-</td>
<td>2013</td>
<td>[31]</td>
</tr>
<tr>
<td>Busulfan</td>
<td>Haematopoietic progenitor cell</td>
<td>Decision tree model using adjusted trial results to compare the treatment and comparator</td>
<td>Yes</td>
<td>Trial</td>
<td>-</td>
<td>2012</td>
</tr>
<tr>
<td>Temozolomide</td>
<td>Anaplastic astrocytoma</td>
<td>Compares trial results to a meta-analysis of trials of the comparator treatment</td>
<td>Yes</td>
<td>Meta-analysis</td>
<td>-</td>
<td>2000</td>
</tr>
<tr>
<td>Temporfin</td>
<td>Head and neck cancer</td>
<td>Trial results for the treatment compared naively to the control arm of a clinical trial for another treatment</td>
<td>Yes</td>
<td>Trial</td>
<td>-</td>
<td>2004, 2005</td>
</tr>
<tr>
<td>Gemtuzumab</td>
<td>Acute myeloid leukaemia</td>
<td>Trial data for the drug compared with hospital data for the comparator. The group were matched for demographics</td>
<td>Yes</td>
<td>Case series</td>
<td>Yes</td>
<td>2002</td>
</tr>
<tr>
<td>Alemtuzumab</td>
<td>Chronic lymphocytic leukaemia</td>
<td>Assumption made of identical efficacy between products, with naive comparison of trials used as sensitivity analysis</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2007</td>
</tr>
<tr>
<td>Imatinib Mesylate</td>
<td>Chronic myeloid leukaemia</td>
<td>Single arm study used for treatment results, which are compared (adjusted) to the results of the comparator taken from an RCT</td>
<td>Yes</td>
<td>Trial</td>
<td>-</td>
<td>2002-2004</td>
</tr>
<tr>
<td>Imatinib Mesylate</td>
<td>Chronic myeloid leukaemia</td>
<td>Single arm study used for treatment results, which are compared to expert opinion for the comparator</td>
<td>Yes</td>
<td>Expert opinion</td>
<td>-</td>
<td>2002-2003</td>
</tr>
<tr>
<td>Imatinib Mesylate</td>
<td>Gastrointestinal stromal tumours</td>
<td>Exponential curve fitted to the trial data for imatinib, naively compared to a historical control from a trial</td>
<td>Yes</td>
<td>Trial</td>
<td>-</td>
<td>2003-2007</td>
</tr>
<tr>
<td>Drug</td>
<td>Disease</td>
<td>Methodology</td>
<td>Comparator Methodology</td>
<td>Year</td>
<td>References</td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>--------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>--------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>Imatinib Mesylate</td>
<td>Gastrointestinal stromal tumours</td>
<td>Trial data for imatinib compared to a case series for the comparator</td>
<td>-</td>
<td>2003</td>
<td>[45]</td>
<td></td>
</tr>
<tr>
<td>Cetuximab</td>
<td>Colorectal cancer</td>
<td>Trial data used for cetuximab, with the comparator effectiveness estimated by reducing the survival based on results from another trial</td>
<td>-</td>
<td>2007</td>
<td>[46,47]</td>
<td></td>
</tr>
<tr>
<td>Cetuximab</td>
<td>Colorectal cancer</td>
<td>Trial data used for cetuximab, using a naive comparison to results from another trial which included the comparator</td>
<td>Yes</td>
<td>2007</td>
<td>[20,48]</td>
<td></td>
</tr>
<tr>
<td>Cetuximab</td>
<td>Colorectal cancer</td>
<td>Trial data used for cetuximab, with threshold analysis performed how ineffective the comparator would need to make cetuximab cost-effective</td>
<td>-</td>
<td>2007</td>
<td>[20,48]</td>
<td></td>
</tr>
<tr>
<td>Cetuximab</td>
<td>Colorectal cancer</td>
<td>Trial data used for cetuximab, with case series data taken from the same hospitals used as the control arm</td>
<td>Yes</td>
<td>2007</td>
<td>[11]</td>
<td></td>
</tr>
<tr>
<td>Trabectedin</td>
<td>Soft tissue sarcoma</td>
<td>Uncontrolled trial data compared naively to pooled data from two trials for the comparator (ifosfamide)</td>
<td>Yes</td>
<td>2008-2011</td>
<td>[49–52]</td>
<td></td>
</tr>
<tr>
<td>Trabectedin</td>
<td>Soft tissue sarcoma</td>
<td>Uncontrolled trial data compared with pooled data from four trials for the comparator, with adjustments made to account for differences in baseline characteristics</td>
<td>Yes</td>
<td>2009-2013</td>
<td>[53–57]</td>
<td></td>
</tr>
<tr>
<td>Trabectedin</td>
<td>Soft tissue sarcoma</td>
<td>Progression free survival taken from studies of the comparators then compared naively. All patients then assumed to have equal post progression survival</td>
<td>Yes</td>
<td>2014</td>
<td>[58,59]</td>
<td></td>
</tr>
<tr>
<td>Cladribine</td>
<td>Hairy cell leukaemia</td>
<td>Naive historical control of results pooled trial data clinical trials. Results from the studies were adjusted for patient disease status, but not demographic characteristics</td>
<td>Yes</td>
<td>2007</td>
<td>[60]</td>
<td></td>
</tr>
<tr>
<td>Gefitinib</td>
<td>Non-small cell lung cancer</td>
<td>Naive comparison of gefitinib uncontrolled clinical study to the best supportive care arm from a RCT in the same disease area</td>
<td>Yes</td>
<td>2004</td>
<td>[61]</td>
<td></td>
</tr>
<tr>
<td>Clofarabine</td>
<td>Acute lymphoblastic leukaemia</td>
<td>Compared outcomes from the drug to results from a registry matched to patient characteristics</td>
<td>Yes</td>
<td>2006, 2007</td>
<td>[62,63]</td>
<td></td>
</tr>
<tr>
<td>Nelorabine</td>
<td>T-cell acute lymphoblastic leukaemia / lymphoma</td>
<td>Trial data for Nelorabine compared to historical trial data for the comparator</td>
<td>Yes</td>
<td>2008, 2009</td>
<td>[64,65]</td>
<td></td>
</tr>
<tr>
<td>Betaine Anhydrous</td>
<td>Homocystinuria</td>
<td>Decision tree comparing results of an uncontrolled study to results seen in a registry</td>
<td>Yes</td>
<td>2010</td>
<td>[66]</td>
<td></td>
</tr>
<tr>
<td>Algglucosidase Alfa</td>
<td>Pompe disease</td>
<td>Uncontrolled data compared to matched registry data</td>
<td>Yes</td>
<td>2007, 2012</td>
<td>[67,68]</td>
<td></td>
</tr>
<tr>
<td>Algglucosidase Alfa</td>
<td>Pompe disease</td>
<td>Trial data compared naively to registry data from the disease area</td>
<td>Yes</td>
<td>2014</td>
<td>[69]</td>
<td></td>
</tr>
<tr>
<td>Drug</td>
<td>Disease/Diagnosis</td>
<td>Data Source Details</td>
<td>Use of Expert Opinion</td>
<td>Comparator Data Details</td>
<td>Yes/No</td>
<td>Case Series</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>----------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>--------</td>
<td>-------------</td>
</tr>
<tr>
<td>Sunitinib Malate</td>
<td>Renal cell carcinoma</td>
<td>Uncontrolled trial data compared to a published case series and a Medicare case series</td>
<td>Yes</td>
<td>Case series</td>
<td>Yes</td>
<td>-</td>
</tr>
<tr>
<td>Sunitinib Malate</td>
<td>Renal cell carcinoma</td>
<td>Expert opinion used to estimate the benefit of sunitinib</td>
<td>Yes</td>
<td>Expert opinion</td>
<td>Yes</td>
<td>-</td>
</tr>
<tr>
<td>Sunitinib Malate</td>
<td>Renal cell carcinoma</td>
<td>Trial data for sunitinib compared to a Finnish case series</td>
<td>Yes</td>
<td>Case series</td>
<td>Yes</td>
<td>-</td>
</tr>
<tr>
<td>Dasatinib</td>
<td>Chronic myeloid leukaemia</td>
<td>Uncontrolled study data for dasatinib compared to trial data for the comparator</td>
<td>Yes</td>
<td>Trial</td>
<td>Yes</td>
<td>-</td>
</tr>
<tr>
<td>Dasatinib</td>
<td>Chronic myeloid leukaemia</td>
<td>Uncontrolled study data for dasatinib compared to trial data for the comparator</td>
<td>Yes</td>
<td>Trial</td>
<td>Yes</td>
<td>-</td>
</tr>
<tr>
<td>Dasatinib</td>
<td>Philadelphia chromosome-positive acute lymphoblastic leukaemia</td>
<td>Historical control for the comparator compared to trial data for dasatinib</td>
<td>Yes</td>
<td>Trial</td>
<td>Yes</td>
<td>-</td>
</tr>
<tr>
<td>Bortezomib</td>
<td>Mantle cell lymphoma</td>
<td>Trial data compared to historical control taken from a Canadian cancer registry</td>
<td>Yes</td>
<td>Registry</td>
<td>Yes</td>
<td>-</td>
</tr>
<tr>
<td>Nilotinib Hydrochloride Monohydrate</td>
<td>Chronic myeloid leukaemia</td>
<td>Results from clinical trials for each of the comparators used in the model without adjustment</td>
<td>Yes</td>
<td>Trial</td>
<td>Yes</td>
<td>-</td>
</tr>
<tr>
<td>Nilotinib Hydrochloride Monohydrate</td>
<td>Chronic myeloid leukaemia</td>
<td>Uncontrolled study results for each of the comparators</td>
<td>Yes</td>
<td>Trial</td>
<td>Yes</td>
<td>-</td>
</tr>
<tr>
<td>Nilotinib Hydrochloride Monohydrate</td>
<td>Chronic myeloid leukaemia</td>
<td>Response rates taken from clinical studies for each of the comparators, then used to predict survival based on this surrogate outcome</td>
<td>Yes</td>
<td>Trial</td>
<td>Yes</td>
<td>-</td>
</tr>
<tr>
<td>Tocofersolan</td>
<td>Vitamin E deficiency due to cholestasis</td>
<td>Trial results for the treatment, with expert opinion used for the comparator</td>
<td>Yes</td>
<td>Expert opinion</td>
<td>Yes</td>
<td>-</td>
</tr>
<tr>
<td>Ofatumumab</td>
<td>Chronic lymphocytic leukaemia</td>
<td>Non-responders assumed to represent the outcome for untreated patients</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Carglumic acid</td>
<td>Chronic hyperammonemia</td>
<td>Comparator data taken from an Italian hospital</td>
<td>Yes</td>
<td>Case series</td>
<td>Yes</td>
<td>-</td>
</tr>
<tr>
<td>Asparaginase Erwinia Chrysanthemi</td>
<td>Acute lymphoblastic leukaemia</td>
<td>Assumption made of equal efficacy for all treatments (cost-minimisation)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Brentuximab Vedotin</td>
<td>Hodgkin's lymphoma</td>
<td>Trial data used for the drug, with a systematic review conducted for comparator data, which was then adjusted for differences in patient characteristics between trials</td>
<td>Yes</td>
<td>Trial</td>
<td>Yes</td>
<td>-</td>
</tr>
<tr>
<td>Defibrotide</td>
<td>Veno-occlusive disease</td>
<td>Trial data used for defibrotide, with then selected patients from a case series used for comparator data</td>
<td>Yes</td>
<td>Historical control</td>
<td>Yes</td>
<td>-</td>
</tr>
<tr>
<td>Drug</td>
<td>Condition</td>
<td>Outcome Measure</td>
<td>Methodological Source</td>
<td>Field</td>
<td>Year</td>
<td>Status</td>
</tr>
<tr>
<td>------------</td>
<td>----------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>-----------------------</td>
<td>-------</td>
<td>------</td>
<td>--------</td>
</tr>
<tr>
<td>Bosutinib</td>
<td>Chronic myeloid leukaemia</td>
<td>Surrogate outcome of response rate taken from each comparator and used to estimate outcomes</td>
<td>Yes</td>
<td>Trial</td>
<td>-</td>
<td>2013</td>
</tr>
<tr>
<td>Bosutinib</td>
<td>Chronic myeloid leukaemia</td>
<td>Time on treatment for each drug assumed to be additive to estimate overall survival</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2013</td>
</tr>
<tr>
<td>Bosutinib</td>
<td>Chronic myeloid leukaemia</td>
<td>Trial data for the drug compared to expert opinion of survival for untreated patients</td>
<td>Yes</td>
<td>Expert opinion</td>
<td>-</td>
<td>2013</td>
</tr>
<tr>
<td>Bosutinib</td>
<td>Chronic myeloid leukaemia</td>
<td>Trial data for the drug compared to expert opinion of survival for untreated patients</td>
<td>Yes</td>
<td>Expert opinion</td>
<td>-</td>
<td>2013</td>
</tr>
</tbody>
</table>
Figure 1: PubMed search terms for cost-effectiveness papers

1. Generic drug name
2. Drug brand name EU
3. Drug brand name US
4. Or 1-3
5. Cost-Benefit Analysis
6. Cost-utility
7. Cost-effectiveness
8. Pharmacoeconomic*
9. health economic*
10. cea
11. cua
12. markov*
13. “patient level simulation”
14. “discrete event simulation”
15. “monte carlo”
16. “decision tree”
17. “quality adjusted life”
18. qaly*
19. qalo*
20. qals
21. “disability adjusted life”
22. hta
23. “health technology assessment”
24. or 5-23
25. 4 AND 24
Figure 2: PRISMA diagram of economic evaluations retrieved from PubMed

- Initial PubMed hits from 74 literature searches: Total = 1202
- 56 Full articles retrieved for review
- Papers excluded (n=1179): Different indication (n=655), Different drug or intervention (n=50), Clinical papers or commentaries (n=341), Not models (n=94), Model based on RCT data (n=39)
- 56 papers reviewed
- Papers excluded (n=27): Different indication (n=7), Different drug or intervention (n=3), Clinical paper or commentary (n=1), Not an economic model (n=6), Model based on RCT data (n=9), Insufficient information reported (n=4)
- 29 papers included
Figure 3: Taxonomy of economic modelling approaches used for estimating incremental benefit from uncontrolled clinical studies

Economic modelling approaches identified:
Total = 51

- Historical control
  Total = 43
- Patients as their own control
  Total = 3
- Cumulative method
  Total = 1
- Threshold analysis
  Total = 1
- Cost minimisation
  Total = 3

Source of comparison data:
- Clinical trial
  Total = 17
- Meta-analysis or pooling of trials
  Total = 5
- Registry or case series
  Total = 14
- Expert opinion
  Total = 7