Relative value assessment: characterizing the benefit of oncology therapies through diverse survival metrics from a US perspective

Richard Macaulay¹
Amit Ahuja²
Ebenezer Ademisoye¹
Ariadna Juarez-Garcia³
James W Shaw⁴

¹PAREXEL Access Consulting, London, UK; ²PAREXEL Access Consulting, Union Territory of Chandigarh, India; ³Bristol-Myers Squibb, Mexico City, Mexico; ⁴Bristol-Myers Squibb, Princeton, NJ, USA

Objectives: The introduction of innovative, high-cost oncology treatments, coupled with mounting budgetary pressures, necessitates value trade-offs across cancer types. Defining value is critical to informing decision-making. A cost-value analysis tool was used to assess relative clinical value from a US perspective using multiple outcome metrics for a variety of metastatic cancers.

Methods: Literature published (January 1, 2000–August 31, 2016) was reviewed to identify outcome metrics for approved treatments for metastatic cancers. Data were extracted or derived for median and mean overall survival (OS), landmark survival rates, and other survival metrics, and compared across treatments vs their respective trial comparators, with and without considering costs.

Results: Reported survival metrics varied by agent within cancer type. For treatment of prostate cancer, abiraterone yielded the highest improvement in 1-year survival rate (13.7%, previously treated), whereas enzalutamide yielded the highest median OS improvement (4.8 months, previously treated) and sipuleucel-T, the highest mean OS improvement (3.6 months, previously untreated) vs their respective trial comparators. For treatment of non-small cell lung cancer vs their respective trial comparators, nivolumab yielded the highest improvement in mean OS (11.9 months) and 3-year survival rate (12.6%), each in previously treated squamous disease, whereas afatinib yielded the highest median OS improvement (4.1 months, previously untreated) EGFR del19 and L858R mutants. Cost-value analysis results varied with the applied survival metric.

Conclusions: Although median OS is the traditional gold standard oncology efficacy metric, it fails to capture long-term survival benefits—the ultimate goal of cancer treatment—offered by new treatment modalities. Diverse metrics are needed for comprehensive value assessments of cancer therapies.

Keywords: value framework, value assessment, immuno-oncology, cost-value analysis

Plain language summary
The efficacy of a cancer therapy is traditionally measured by median overall survival (OS; the point in time at which half of the patients receiving therapy have died). However, newer therapies such as immuno-oncology agents have shown long-term survival benefits in some patients that are not adequately captured by median OS. We reviewed survival outcomes in pivotal clinical trials for therapies approved for metastatic cancers in the US between January 1, 2000 and August 31, 2016.

Improvements in median OS, average OS and landmark survival (the proportion of patients alive after receiving therapy for a specific time) with approved therapies were assessed relative to trial comparator arms. Relative clinical value was considered by applying the total cost of therapy per patient to the relative survival improvements.
Survival advantage and relative value of approved agents differed depending on the survival metric assessed. Median OS may undervalue treatments that provide long-term survival benefit, and average OS and landmark survival may provide more accurate measures of such long-term survival. When assessing the value of oncology therapies, decision-makers should consider a diverse range of oncology survival metrics.

Introduction

Substantial unmet needs remain in cancer, and long-term survival continues to be an elusive goal for most patients with advanced cancer. Further complicating treatment goals are the continually rising costs associated with cancer care, which are outpacing other types of healthcare costs. A variety of factors are driving the budgetary impact of cancer care, including increasing incidence with a growing and aging population, a shift toward more costly personalized care, increased cancer survivorship, and increasing availability and adoption of new branded therapies.

Against this backdrop of increasing cancer incidence and healthcare costs, many healthcare budget holders are facing mounting pressure to control costs while funding optimal care for a growing population. Value trade-offs across cancer types are, thus, becoming increasingly important in resource allocation decision-making. Defining value is critical in determining what treatments should be reimbursed. Each stakeholder group (patients, caregivers, physicians, and payers) defines the value of cancer treatment differently, resulting in an ongoing debate and a lack of agreement regarding the meaning of value. Oncology treatment guidelines generally do not account for treatment costs. Nevertheless, multiple traditionally cost-agnostic provider organizations have created “scorecard”-based frameworks to define the value of oncology treatments, including the American Society of Clinical Oncology (ASCO), the European Society for Medical Oncology (ESMO), and the Memorial Sloan Kettering Cancer Center (MSKCC). As of 2015, National Comprehensive Cancer Network (NCCN) guidelines have included Evidence Blocks, a visual representation for each agent or regimen of five key value measures, one of which is affordability. The Institute for Clinical and Economic Review has also incorporated economic analyses, but, like the NCCN Evidence Blocks, these use an expert panel-driven rather than a scorecard approach. These initiatives demonstrate increasing concern over escalating healthcare costs.

One challenge in assessing value is identifying the appropriate efficacy metrics. Traditional “gold-standard” efficacy metrics such as median overall survival (OS) or progression-free survival (PFS) focus on a single point, where 50% of the population has had an event. Thus, they are not well suited to reflect the long-term survival benefits of new treatment modalities, such as immuno-oncology agents, which demonstrate prolonged survival in a proportion of patients. Despite this, several common value frameworks utilize median OS as a key driver of value scoring. To capture long-term survival benefits, and thus assess the full value of newer treatment modalities, there is a need to report and evaluate nontraditional efficacy endpoints such as landmark and mean OS.

The objectives of this study were to: 1) compare the value of approved oncology agents, in particular new immuno-oncology agents, as assessed by a variety of efficacy metrics vs median OS, within and between tumor types and 2) apply a recently developed relative value assessment (RVA) tool for conducting both outcomes assessments and cost-value analyses. These analyses were conducted from a US perspective. Application of the RVA tool to clinical trial data may be useful in assessing managed care decision-making.

Methods

This analysis included any therapy approved by the US Food and Drug Administration (FDA) between January 1, 2000 and August 31, 2016, inclusive, for the treatment of stage III/IV of the following cancers: breast cancer, bladder cancer, colorectal cancer (CRC), lymphomas, gastric cancer, glioblastoma multiforme, hepatocellular carcinoma, myelodysplastic syndromes, melanoma, multiple myeloma, non-small cell lung cancer (NSCLC), prostate cancer, renal cell carcinoma (RCC), small cell lung cancer, and squamous cell carcinoma of the head and neck. Single-arm trials were excluded, as survival metrics were evaluated for each drug vs its direct trial comparator. Two kinds of subanalyses were performed: within-tumor analyses and pan-tumor analyses. The within-tumor analyses focused on NSCLC and prostate cancer, as these reflect common cancers, for which baseline survival rates are among the lowest and highest, respectively, across all cancers. The pan-tumor analyses included a selection of tumor types identified as the most common causes of cancer-related deaths in the US (breast cancer, CRC, melanoma, NSCLC, and RCC).

Survival metrics

For each agent and its indicated tumor type (prostate, NSCLC, breast cancer, melanoma, RCC, and CRC), survival
metrics were obtained from FDA product labels and European Medicines Agency (EMA) Public Assessment Reports, or relevant clinical trial publications in cases where these were not in the FDA label or EMA Public Assessment Report (see Table S1). Kaplan–Meier (KM) OS curves were extracted and digitized using the Engauge software package (http://markummitchell.github.io/engauge-digitizer/), allowing for the calculation of mean OS (calculated as the area under the KM curve) and landmark survival at 1, 2, and 3 years for each agent and tumor type. After extracting survival metrics, the difference between an intervention and its trial comparator was calculated.

A confounding factor in the calculation of mean OS is the maturity of data cutoff, which may particularly disadvantage newer agents that have less mature OS data available, particularly those that provide long-term OS benefits. To address this, KM OS curves were digitized, and reconstructed individual patient-level data were generated based on the code from Guyot et al.21 Parametric distributions, including spline models, were fitted to the reconstructed individual patient-level data to provide extrapolation beyond reported cutoffs at 15 years. The parametric curve with the best fit was determined by goodness-of-fit statistics; specifically, those with the lowest Akaike information criterion were chosen and then the visual best fit was used to validate this selection. A 15-year cutoff was used for all extrapolated curves; an assumption was that at the end of year 15 survival would be 0%. Mean OS and landmark survival were recalculated using this extrapolated curve set for therapies in melanoma, RCC, and NSCLC (see Table S2), which were the first three approved indications of immuno-oncology agents.

Cost-value analyses

In addition to calculating survival metrics, relative clinical value across tumor types (breast cancer, melanoma, NSCLC, RCC, and CRC) was assessed in cost-value analyses using the RVA tool by plotting total drug cost vs a given survival outcome. Calculation of total drug costs was based on monthly costs multiplied by therapy duration, as noted by data in the product label, determined based on median duration of administration, median PFS, or median time to progression, ranked in descending order of preference, as available. Loading dose for applicable agents, as indicated by the product label, was also incorporated into total cost. Total drug costs were obtained from the IHS Market Life Sciences PharmOnline International database.22 All costs are expressed in US dollars.

For all therapies with available data, the relative value of each product for a given clinical metric, the relative value of each product for a given clinical metric, the relative value of each product for a given clinical metric, the relative value of each product for a given clinical metric, the relative value of each product for a given clinical metric. The cost of avoiding death at 1 year was calculated by multiplying the number needed to treat to avoid a death at 1 year by the per-patient treatment cost. This number was inverted to calculate the number of additional patients alive at 1 year per dollar spent. The agent with the lowest value was set equal to one and used to calculate the relative additional persons alive at 1 year for all other agents.

Results

Survival metrics

In the individual tumor analyses, efficacy results for individual agents varied by survival metric. For treatment of prostate cancer, for which all metrics were based on reported KM OS curve data, abiraterone yielded the highest improvement in 1-year survival rate (13.7%, previously treated, Figure 1A), whereas sipuleucel-T yielded the highest median OS improvement (4.8 months, previously untreated, Figure 1B) and enzalutamide yielded the highest mean OS improvement (3.6 months, previously treated, Figure 1C) among all agents vs their respective trial comparators.

For treatment of NSCLC, nivolumab yielded the highest improvement in 1-year survival rate (18%, previously treated squamous disease) and mean OS (5.0 months, previously treated squamous disease) among all agents vs their respective trial comparators (figures not presented), whereas afatinib yielded the highest improvement in median OS (4.1 months, previously untreated EGFR del19 and L858R mutants, Figure 2A) based on reported KM curves. In the extrapolated analysis, which helps to account for differences in data maturity, nivolumab again yielded the highest improvement in 3-year survival rate (12.6%, previously treated squamous disease,
In the case of immuno-oncology agents used to treat NSCLC (nivolumab in previously treated disease, irrespective of programmed death ligand 1 [PD-L1] expression and pembrolizumab in previously treated ≥ 1% PD-L1-positive disease; see Table S1), the greatest survival benefits vs their respective trial comparators were apparent when mean OS

Figure 2B) and in mean OS (11.8 months, previously treated squamous disease, Figure 2C).
Figure 2 Non-small cell lung cancer survival improvement. (A) Improvement in median OS based on reported Kaplan–Meier OS curves, (B) improvement in 3-year OS, and (C) improvement in mean OS for each agent vs its respective trial comparator, based on fitted Kaplan–Meier OS curves that extrapolate survival beyond the reported cutoffs; excludes interventions where relevant Kaplan–Meier OS curves were not identified (ie, afatinib, nintedanib, and pemetrexed [2L]). Any drug compared with placebo or best supportive care (offers a lower clinical benchmark against which it is easier to demonstrate relative value) was excluded (ie, pemetrexed [maintenance], docetaxel, and erlotinib [2/3L]).

**Abbreviations:** 1L, first line; 2L, second line; 3L, third line; Af, afatinib; Bev, bevacizumab; Criz, crizotinib; Erlot, erlotinib; Gefit, gefitinib; Nab-pac, nab-paclitaxel; Neci, necitumumab; Nivo, nivolumab; NSQ, nonsquamous; OS, overall survival; PD-L1, programmed death ligand 1; Pemet, pemetrexed; Pembro, pembrolizumab 2 mg/kg; Ramu, ramucirumab; SQ, squamous.
and 3-year survival rate improvements (based on extrapolated curves) were used as the comparative metrics (Figures 2B and C). By comparison, when median OS improvement based on reported curves was used to compare agents (Figure 2A), the benefits of immuno-oncology drugs vs their respective trial comparators were comparable with those of many targeted alternatives in NSCLC. Furthermore, the magnitude of variation among NSCLC agents across the different survival metrics was greater than that observed in prostate cancer, where immuno-oncology agents were not used.

Cost-value analyses
Results of the pan-tumor cost-value analyses are shown in Figures 3–6. Presentation of these data as a single-variable plot, in terms of the relative number of additional patients alive at 1 year per US dollar spent on a range of treatments for NSCLC, is provided in Figure S1.

Therapies used for tumors and lines of therapy with higher baseline median survival were generally associated with the greatest absolute median survival gain compared with their trial comparators; that is, in these settings, the same relative risk reduction would result in a larger absolute gain in median OS (eg, first-line melanoma and breast cancer).23 Using 1-year survival as the metric, some of the agents with greatest benefit vs their trial comparators were for NSCLC, where baseline survival is very low.23 However, two (docetaxel and pemetrexed) of the three NSCLC agents with the greatest survival gain vs their respective comparators (docetaxel, pemetrexed, and nivolumab) had best supportive care as their trial comparator, providing a low clinical benchmark against which greater efficacy gains can be demonstrated. Similarly, based on the cost-value analyses, nivolumab + ipilimumab appeared less favorable than nivolumab monotherapy in first-line melanoma in terms of the mean OS benefit over their respective trial comparators. The comparator arm for nivolumab + ipilimumab in CheckMate 069 was ipilimumab,24 which is a higher clinical benchmark than cytotoxic chemotherapy (dacarbazine, the comparator for nivolumab monotherapy in CheckMate 066),23 making demonstration of relative value more challenging. As with this example, branded combination therapies may be rated as having a lower relative clinical value than their constituent monotherapies. Thus, the relative value of branded combination agents can appear comparable with or even lower than that offered by their constituent monotherapies; therefore, it may be more appropriate to evaluate the value of branded combination regimens separately (Figure 4).

Results varied with the applied survival metric, with some agents achieving a higher cost-value for some metrics, but not others, as can be seen by differences in relative value from 1-year survival rates based on reported KM curves (Figures 3 and 4), median OS based on reported KM curves (Figure 5), and mean OS based on extrapolated curves (Figure 6). For example, the intervention with the greatest increase in median OS was pertuzumab + trastuzumab for breast cancer, at a cost that appeared to provide above-average efficacy benefits compared with other available therapies (as indicated by its relative position over the regression line in Figure 5). In contrast, when looking across agents at increases in 1-year survival (Figure 4), pertuzumab + trastuzumab for breast cancer provided a much less favorable efficacy profile (5.2% increase in 1-year survival), and at a cost (more than $196,000 per patient) far in excess of the average value offered (as indicated by its position relative to the regression line). This discrepancy may illustrate limitations in median OS as an efficacy metric, as it cannot take into account any long-term survival benefits after the median has been reached, unlike mean OS and landmark survival analyses.

Discussion
This analysis demonstrated that therapy benefits differed depending on the survival metric used, with evaluations of oncology therapies based on improvements in median OS yielding some very distinct results vs other survival metrics. Although median OS is among the most commonly used metrics in oncology trials, it fails to take into account long-term survival benefits after the median has been reached and, thus, is not well suited to assess the value of immuno-oncology therapies, for which efficacy is typified by KM survival curves showing delayed but sustained separation.

Alternative survival metrics include mean OS and landmark survival rates. In the pan-tumor analysis, some differences in cost-value patterns were also observed with the 1-year survival metric (Figure 3) compared with the mean OS metric (Figure 6). Although median OS provides an estimate that cannot be biased by a small number of outlying cases, it calculates survival for the whole cohort based on only the first 50% of patients who do not survive.26,27 Mean OS offers a more comprehensive assessment of long-term survival benefits, as it captures the complete survival curve, including the “tail”. In terms of value for money, many healthcare payers and health technology assessment agencies consider mean OS as the standard metric for cancer treatment cost-effectiveness analyses.27 However, a limitation of assessing value using mean OS calculations based on the study duration
Deemed appropriate by regulatory bodies for each treatment’s Phase III pivotal trial is that duration is not consistent among trials, which may confound comparisons. Although mean OS offers a more comprehensive assessment of long-term survival benefits, using this metric requires extrapolation of KM OS curves beyond reported cutoffs. Extrapolation in this study was undertaken for therapies indicated for NSCLC, RCC, or melanoma (the first three cancers with an approved immuno-oncology agent). Use of digitized KM curves, rather than actual data points, where patient-level data were not available, is a limitation of our analysis. Furthermore, parametric curve selection, particularly with an immature data set, can be subject to interpretation, as distinct curves with similar goodness-of-fit statistics can introduce variability in outcomes.

Landmark survival rates can be highly effective for comparing therapies by enabling survival comparisons after the median has been reached. However, data maturity and cutoffs varied among trials included in this assessment. As selection of a later time point for landmark survival rates could severely constrain the comparator set, 1-year survival rates based on reported KM OS curves (where available) were used. Using such an early time point, however, may not account for long-term survival benefits of selected therapies, and indeed may even precede the median OS cutoff for some tumor types. Utilizing fitted parametric KM curves to provide survival data beyond the reported cutoffs and the selection of landmark survival at a later time point for landmark survival rates could severely constrain the comparator set, 1-year survival rates based on reported KM OS curves (where available) were used. Using such an early time point, however, may not account for long-term survival benefits of selected therapies, and indeed may even precede the median OS cutoff for some tumor types. Utilizing fitted parametric KM curves to provide survival data beyond the reported cutoffs and the selection of landmark survival at a later time point can provide a consistent set of more mature landmark survival readouts, but is subject to limitations, as discussed above.

In addition to survival, health-related quality-of-life (HRQoL) is an important metric that provides detailed information from the patient perspective, and is increasingly being reported in clinical studies. However, unlike survival
outcomes, HRQoL is not reported consistently across studies. Furthermore, diversity in HRQoL metrics used across trials, variability in timing of PRO data reporting, and the percentage of patients missing PRO data limit comparability within and across tumor types. Therefore, HRQoL metrics were not included in this analysis.

Our results showed variation in the relative value of the agents assessed, particularly between median OS and other metrics. Nevertheless, the findings showed no clear ideal metric, with variations observed in the relative value of treatments, even between landmark survival rates and mean OS metrics. Moreover, a higher relative survival outcome on a given metric did not necessarily translate into greater economic value. These discrepancies emphasize the need to incorporate a wide range of metrics into assessments of the efficacy benefits of different agents in a pan-tumor as well as in individual tumor comparison of value, or to be able to appropriately target value metrics to the agents of interest.

Pan-tumor value comparisons should be interpreted with caution owing to heterogeneity between trials (particularly variations in comparators, but also in populations, study design, outcomes, and timing of assessments). Furthermore, baseline survival expectations differ among tumor types and lines of therapy and must be considered when evaluating relative value so as not to discriminate unfairly against agents that treat cancers with the highest unmet need. Although the ASCO Value in Cancer Care Task Force has expressed concern about cross-trial comparisons due to potential biases, the task force has also suggested that the ASCO value framework tool could be used to facilitate individual patient treatment decision-making based on data from multiple trials for a specific indication.

Another factor that may affect relative value is variation in patient crossover rates from the control to the active agent arm, which can confound OS results. For example, in Figure 4, the two regimens with the greatest survival benefits (lenvatinib + everolimus and cobimetinib + vemurafenib) were from trials where crossover was not allowed, whereas the regimen with the next highest survival benefit (nivolumab + ipilimumab) was from a trial where more than 55% of patients...
An additional limitation of RVA is that it involves naïve trial comparisons, in distinct cancer types and lines of therapy, and in different patient populations. This includes making comparisons of therapies approved in biomarker-selected vs nonbiomarker selected populations. For example, in Figure 5, the anti-PD-1 agent pembrolizumab appears to produce a survival benefit in previously treated NSCLC in the pembrolizumab arm crossed over to receive nivolumab plus ipilimumab. Pembrolizumab, which also had ipilimumab as its trial comparator for first-line and second-line treatment of melanoma, appeared to offer a mean survival benefit vs ipilimumab of a similar magnitude to that offered by nivolumab plus ipilimumab vs ipilimumab for first-line treatment of melanoma; however, crossover was not permitted in the pembrolizumab trial.

An additional limitation of RVA is that it involves naïve trial comparisons, in distinct cancer types and lines of therapy, and in different patient populations. This includes making comparisons of therapies approved in biomarker-selected vs nonbiomarker selected populations. For example, in Figure 5, the anti-PD-1 agent pembrolizumab appears to produce a survival benefit in previously treated NSCLC that is within the range of the anti-PD-1 agent nivolumab in previously treated squamous and non-squamous NSCLC, but at a lower cost. However, as of the analysis end date, pembrolizumab was only approved in PD-L1+ patients in this line of therapy, whereas nivolumab was approved in an all-comers population. As an alternative, more complex meta-analytic approaches (if feasible given available data resources), instead of this type of pan-tumor value assessment tool, could be used to reduce bias. The results of the individual tumor analyses for NSCLC and prostate cancer can address some of these limitations to an extent, and provide an alternate means for assessing the relative value assessment approach.

Another limitation in this cost-value analysis is that each therapy was assigned a cost based on reported monthly cost, multiplied by therapy duration (median duration of administration, median PFS, or median time to progression, based on available data), which varied between trials. Although costs could be underestimated by using median instead of mean therapy duration, this bias was assumed to be relatively consistent across agents because similar methods...
were used to estimate costs for all comparators. Furthermore, assigned cost-value only incorporated drug cost as adjusted for duration and loading dose, and did not take into account symptom burden, treatment safety profile, HRQoL, healthcare resource utilization, or use of postprogression therapies. These factors are particularly important with newer, frequently less toxic therapies compared with traditional chemotherapy. Additionally, the cost-value analysis did not take into account drug cost discounting, which might be available among certain payers, or compassionate use programs or patient rebates. As with any complex analysis that entails input of a large and diverse body of evidence, only FDA-approved drugs during the study period (January 1, 2000—August 21, 2016) are included; later approvals were not included in the analysis. Finally, the predefined inclusion criteria may have resulted in the evaluated agents not being representative of the entire treatment landscape, thus potentially introducing selection bias.

Despite the limitations noted here, from which no value assessment tool is exempt, the RVA tool utilized in this study provides a framework for comparing oncology therapies both within and across tumor types. As such, it adds to the range of value tools and metrics available that include treatment costs for valuing oncology treatments.\textsuperscript{12–15,30} US payers often look to NCCN and ASCO clinical guidelines as part of their reimbursement decision-making,\textsuperscript{11} which either do not currently account for treatment costs (ASCO) or provide a five-level affordability rating (NCCN). Several of the existing value scorecards come from provider bodies (ASCO, NCCN, and MSKCC), suggesting that traditionally cost-agnostic stakeholders are becoming increasingly concerned over the escalating costs of care. As the cost of care
continues to rise, payers may be increasingly forced to make value trade-offs, and this tool can be used to support funding prioritization.

While some of these tools base scores heavily on median OS or PFS benefits, the RVA tool incorporates a wide range of efficacy metrics that, unlike median OS, can account for long-term survival benefits. As shown in this study, valuing oncology therapies based on improvements in median OS alone potentially undervalues new treatment modalities associated with long-term survival benefits. Ongoing review and discussion of the range of indices to define and measure cost-effectiveness is needed to reach a consensus, or at least more broadly accepted criteria, for defining value, particularly for novel therapies.

Further research could include a systematic methodological standardization for parametric curve selection for the calculation of mean OS and long-term landmark OS endpoints alongside aligning best value metrics and reconciling when results differ. The RVA approach demonstrated here can compare a wide range of efficacy metrics for treatments across a broad range of cancers that can be utilized to better inform decisions by payers and providers in managed care.

Conclusions
In the era of expanding healthcare costs and budgetary pressure, payers will increasingly focus on areas of high and expanding costs, of which cancer care is prominent. Physicians are also becoming attuned to treatment costs, and it is notable that many of the value frameworks and scorecards are provider-led. Although median OS is the traditional gold standard oncology efficacy metric, it should not be considered in isolation of other metrics, because it does not capture long-term survival benefits, the ultimate goal of cancer treatment. The metric selected to define the value of an oncology therapy must be carefully considered. Value frameworks incorporating diverse metrics, such as the one presented here, are needed for comprehensive value assessments of cancer therapies, particularly for new treatment modalities such as immuno-oncology.

Acknowledgments
Medical writing assistance was provided by Karen Smoyer, PhD, and Beth Burke, PhD, CMPP, of Evidence Scientific Solutions Inc, and was funded by Bristol-Myers Squibb. Bristol-Myers Squibb commissioned and funded this study; participated in the analysis and interpretation of the data; contributed to the preparation, review, and approval of the article for publication; and covered costs for medical writing and editorial assistance.

Disclosure
JWS and AJG report employment by Bristol-Myers Squibb during the conduct of this study; JWS is a stockholder of Bristol-Myers Squibb. RM, AA, and EA are employed by PAREXEL International, the entity paid for analyses in this manuscript. This manuscript includes information presented at the International Society for Pharmacoeconomics and Outcomes Research 21st Annual International Meeting, May 2016, in Washington DC, USA and published as a congress abstract at the following site: https://www.valueinhealthjournal.com/article/S1098-3015(16)01616-8/fulltext.

References
# Supplementary materials

## Table S1 Data included in the survival and cost analyses

<table>
<thead>
<tr>
<th>Study treatment</th>
<th>Histology</th>
<th>Line of therapy</th>
<th>Trial</th>
<th>Comparator</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Melanoma</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dabrafenib</td>
<td>NA</td>
<td>1L</td>
<td>BREAK-3</td>
<td>Dacarbazine</td>
</tr>
<tr>
<td>Dabrafenib + trametinib</td>
<td>NA</td>
<td>1L</td>
<td>COMBI-V</td>
<td>Vemurafenib</td>
</tr>
<tr>
<td>Dabrafenib + trametinib</td>
<td>NA</td>
<td>2L</td>
<td>COMBI-D</td>
<td>Dabrafenib</td>
</tr>
<tr>
<td>Ipilimumab</td>
<td>NA</td>
<td>1L</td>
<td>MDX-010-20</td>
<td>Ipilimumab</td>
</tr>
<tr>
<td>Nab-paclitaxel</td>
<td>NA</td>
<td>1L</td>
<td>CA031</td>
<td>Paclitaxel</td>
</tr>
<tr>
<td>Nivolumab</td>
<td>NA</td>
<td>2L+</td>
<td>CheckMate 037</td>
<td>Dacarbazine or carboplatin + paclitaxel</td>
</tr>
<tr>
<td>Nivolumab</td>
<td>NA</td>
<td>1L</td>
<td>CheckMate 066</td>
<td>Dacarbazine</td>
</tr>
<tr>
<td>Nivolumab + ipilimumab</td>
<td>NA</td>
<td>1L</td>
<td>CheckMate 069</td>
<td>Ipilimumab</td>
</tr>
<tr>
<td>Pembrolizumab</td>
<td>NA</td>
<td>1L/2L</td>
<td>KEYNOTE-006</td>
<td>Ipilimumab</td>
</tr>
<tr>
<td>Pembrolizumab</td>
<td>NA</td>
<td>2L+</td>
<td>KEYNOTE-002</td>
<td>Investigator’s choice chemotherapy</td>
</tr>
<tr>
<td>Trametinib</td>
<td>NA</td>
<td>1L/2L</td>
<td>METRIC</td>
<td>Dacarbazine or paclitaxel</td>
</tr>
<tr>
<td>Vemurafenib</td>
<td>NA</td>
<td>1L</td>
<td>BRIM-3</td>
<td>Dacarbazine</td>
</tr>
<tr>
<td><strong>NSCLC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Afatinib</td>
<td>NSQ</td>
<td>1L</td>
<td>LUX LUNG 3</td>
<td>Cisplatin-based chemotherapy</td>
</tr>
<tr>
<td>Bevacizumab + paclitaxel + carboplatin</td>
<td>NSQ</td>
<td>1L</td>
<td>E4599</td>
<td>Paclitaxel + carboplatin</td>
</tr>
<tr>
<td>Crizotinib</td>
<td>NSQ</td>
<td>2L+</td>
<td>PROFILE 1007</td>
<td>Docetaxel or pemetrexed</td>
</tr>
<tr>
<td>Docetaxel</td>
<td>NSQ</td>
<td>2L+</td>
<td>TAX 320</td>
<td>Pemetrexed or vinorelbine</td>
</tr>
<tr>
<td>Erlotinib</td>
<td>NSQ</td>
<td>1L</td>
<td>ERTAC</td>
<td>Standard chemotherapy</td>
</tr>
<tr>
<td>Erlotinib</td>
<td>NSQ</td>
<td>2L/3L</td>
<td>BR.21</td>
<td>Placebo</td>
</tr>
<tr>
<td>Gefitinib</td>
<td>NSQ</td>
<td>1L</td>
<td>IPASS</td>
<td>Carboplatin + paclitaxel</td>
</tr>
<tr>
<td>Nab-paclitaxel</td>
<td>NSQ</td>
<td>1L</td>
<td>CA031</td>
<td>Solvent-based paclitaxel</td>
</tr>
<tr>
<td>Necitumumab + gemcitabine + cisplatin</td>
<td>SQ</td>
<td>1L</td>
<td>SQUIRE trial</td>
<td>Gemcitabine + cisplatin</td>
</tr>
<tr>
<td>Nivolumab</td>
<td>SQ</td>
<td>2L</td>
<td>CheckMate 017</td>
<td>Docetaxel</td>
</tr>
<tr>
<td>Nivolumab</td>
<td>SQ</td>
<td>2L+</td>
<td>CheckMate 057</td>
<td>Docetaxel</td>
</tr>
<tr>
<td>Pembrolizumab (2 mg/kg, ≥1% PD-L1)</td>
<td>NA</td>
<td>2L+</td>
<td>KEYNOTE-001</td>
<td>Docetaxel</td>
</tr>
<tr>
<td>Pemetrexed</td>
<td>NA</td>
<td>1L</td>
<td>FDA study</td>
<td>Cisplatin + gemcitabine</td>
</tr>
<tr>
<td>Pemetrexed</td>
<td>NA</td>
<td>1L maintenance</td>
<td>PARAMOUNT</td>
<td>Placebo</td>
</tr>
<tr>
<td>Ramucirumab</td>
<td>NA</td>
<td>1L maintenance (switch)</td>
<td>JMEN</td>
<td>Placebo</td>
</tr>
<tr>
<td><strong>Renal cell carcinoma</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Axitinib</td>
<td>NA</td>
<td>2L</td>
<td>AXIS</td>
<td>Sorafenib</td>
</tr>
<tr>
<td>Cabozantinib</td>
<td>NA</td>
<td>2L+</td>
<td>METEOR trial</td>
<td>Everolimus</td>
</tr>
<tr>
<td>Everolimus</td>
<td>NA</td>
<td>2L</td>
<td>RECORD-I</td>
<td>Placebo</td>
</tr>
<tr>
<td>Lenvatinib + everolimus</td>
<td>NA</td>
<td>2L+</td>
<td>Study 2 FDA label</td>
<td>Everolimus</td>
</tr>
<tr>
<td>Nivolumab</td>
<td>NA</td>
<td>2L</td>
<td>CheckMate 025</td>
<td>Everolimus</td>
</tr>
<tr>
<td>Pazopanib</td>
<td>NA</td>
<td>1/2L</td>
<td>VEG105192</td>
<td>Placebo</td>
</tr>
<tr>
<td>Sorafenib</td>
<td>NA</td>
<td>2L</td>
<td>TARGET</td>
<td>Placebo</td>
</tr>
<tr>
<td>Sunitinib</td>
<td>NA</td>
<td>1L</td>
<td>A6181034</td>
<td>IFN-alpha</td>
</tr>
<tr>
<td>Temsirolimus</td>
<td>NA</td>
<td>1L</td>
<td>3066K1-304-WW</td>
<td>IFN-alpha</td>
</tr>
<tr>
<td><strong>Breast cancer</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nab-paclitaxel</td>
<td>NA</td>
<td>2L+</td>
<td>CA012-0</td>
<td>Paclitaxel</td>
</tr>
<tr>
<td>Everolimus</td>
<td>NA</td>
<td>2L+</td>
<td>BOLERO-2</td>
<td>Placebo</td>
</tr>
<tr>
<td>Capecitabine + docetaxel</td>
<td>NA</td>
<td>2L+</td>
<td>Xeloda SOI4999</td>
<td>Docetaxel</td>
</tr>
<tr>
<td>Eribulin mesylate</td>
<td>NA</td>
<td>2L/3L</td>
<td>EMBRACE</td>
<td>Physician’s treatment of choice</td>
</tr>
<tr>
<td>Trastuzumab + docetaxel</td>
<td>NA</td>
<td>3L</td>
<td>Herceptin M77001</td>
<td>Docetaxel</td>
</tr>
<tr>
<td>Ado-trastuzumab emtansine</td>
<td>NA</td>
<td>2L+</td>
<td>EMILIA</td>
<td>Lapatinib + capecitabine</td>
</tr>
<tr>
<td>Pertuzumab + trastuzumab + docetaxel</td>
<td>NA</td>
<td>1L</td>
<td>CLEOPATRA</td>
<td>Placebo + trastuzumab + docetaxel</td>
</tr>
<tr>
<td>Trastuzumab emtansine + placebo</td>
<td>NA</td>
<td>1L</td>
<td>MARIANNE</td>
<td>Trastuzumab + taxane (docetaxel/ paclitaxel)</td>
</tr>
<tr>
<td>Lapatinib</td>
<td>NA</td>
<td>1L</td>
<td>EGF 30008</td>
<td>Letrozole + placebo</td>
</tr>
<tr>
<td>Lapatinib + capecitabine</td>
<td>NA</td>
<td>2L+</td>
<td>EGF100151</td>
<td>Capecitabine</td>
</tr>
</tbody>
</table>

(Continued)
### Table S1 (Continued)

<table>
<thead>
<tr>
<th>Study treatment</th>
<th>Histology</th>
<th>Line of therapy</th>
<th>Trial</th>
<th>Comparator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colorectal cancer</td>
<td>Bevacizumab&lt;sup&gt;112,113&lt;/sup&gt;</td>
<td>NA</td>
<td>1L</td>
<td>NO16966</td>
</tr>
<tr>
<td></td>
<td>Bevacizumab&lt;sup&gt;112,124,125&lt;/sup&gt;</td>
<td>NA</td>
<td>1L</td>
<td>AVF2107</td>
</tr>
<tr>
<td></td>
<td>Capcitabine&lt;sup&gt;95,96,126&lt;/sup&gt;</td>
<td>NA</td>
<td>1L</td>
<td>Xeloda pooled (SO14695, SO14796)</td>
</tr>
<tr>
<td></td>
<td>Cetuximab (RAS wild-type population)&lt;sup&gt;117–129&lt;/sup&gt;</td>
<td>NA</td>
<td>1L</td>
<td>CRYSTAL</td>
</tr>
<tr>
<td></td>
<td>Cetuximab (KRAS wild-type population)&lt;sup&gt;129–131&lt;/sup&gt;</td>
<td>NA</td>
<td>2L+</td>
<td>CORRECT</td>
</tr>
<tr>
<td></td>
<td>Regorafenib&lt;sup&gt;112–113&lt;/sup&gt;</td>
<td>NA</td>
<td>3L+</td>
<td>RECURSE trial (TGU-TAS-301 trial)</td>
</tr>
<tr>
<td></td>
<td>Trifluridine + tipiracil + BSC&lt;sup&gt;115–117&lt;/sup&gt;</td>
<td>NA</td>
<td>1L</td>
<td>PRIME (KRAS wild-type)</td>
</tr>
<tr>
<td></td>
<td>Panitumumab + FOLFOX&lt;sup&gt;4&lt;/sup&gt;&lt;sup&gt;118–141&lt;/sup&gt;</td>
<td>NA</td>
<td>2L</td>
<td>VELOUR</td>
</tr>
<tr>
<td>Prostate cancer</td>
<td>Abiraterone&lt;sup&gt;145–149&lt;/sup&gt;</td>
<td>NA</td>
<td>2L</td>
<td>COU-AA-301</td>
</tr>
<tr>
<td></td>
<td>Abiraterone&lt;sup&gt;145,147,148,150–153&lt;/sup&gt;</td>
<td>NA</td>
<td>1L</td>
<td>COU-AA-302</td>
</tr>
<tr>
<td></td>
<td>Cabazitaxel&lt;sup&gt;154&lt;/sup&gt;</td>
<td>NA</td>
<td>2L</td>
<td>TROPIC</td>
</tr>
<tr>
<td></td>
<td>Enzalutamide&lt;sup&gt;155–159&lt;/sup&gt;</td>
<td>NA</td>
<td>1L</td>
<td>PREVAIL</td>
</tr>
<tr>
<td></td>
<td>Enzalutamide&lt;sup&gt;155–157,159,160&lt;/sup&gt;</td>
<td>NA</td>
<td>2L</td>
<td>AFFIRM</td>
</tr>
<tr>
<td></td>
<td>Radium-223 dichloride&lt;sup&gt;161–163&lt;/sup&gt;</td>
<td>NA</td>
<td>1L</td>
<td>ALSYMPCA</td>
</tr>
<tr>
<td></td>
<td>Sipuleucel-T&lt;sup&gt;164,165&lt;/sup&gt;</td>
<td>NA</td>
<td>1L</td>
<td>IMPACT</td>
</tr>
</tbody>
</table>

**Abbreviations:** 1L, first line; 2L, second line; BSC, best support care; IFN, interferon; NA, not applicable; NSCLC, non-small cell lung cancer; NSQ, non-squamous; PD-L1, programmed death ligand 1; SQ, squamous.

### Table S2 Parametric curves selected for OS extrapolation calculations

<table>
<thead>
<tr>
<th>Trial</th>
<th>Study treatment</th>
<th>Comparator</th>
<th>Parametric curve selected – treatment</th>
<th>Parametric curve selected – comparator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Melanoma</td>
<td>BREAK-3</td>
<td>Dabrafenib</td>
<td>Dacarbazine</td>
<td>Log-normal</td>
</tr>
<tr>
<td></td>
<td>COMBI-V</td>
<td>Dabrafenib + trametinib</td>
<td>Vemurafenib</td>
<td>Generalized gamma</td>
</tr>
<tr>
<td></td>
<td>COMBI-D</td>
<td>Dabrafenib + trametinib</td>
<td>Dabrafenib</td>
<td>Log-logistic</td>
</tr>
<tr>
<td></td>
<td>MDX010-20</td>
<td>Ipilimumab</td>
<td>Gp100</td>
<td>Generalized gamma</td>
</tr>
<tr>
<td></td>
<td>CA031</td>
<td>Nabi-paclitaxel</td>
<td>Paclitaxel</td>
<td>Log-logistic</td>
</tr>
<tr>
<td></td>
<td>CheckMate 037</td>
<td>Nivolumab</td>
<td>Dacarbazine or carboplatin + Paclitaxel</td>
<td>Log-normal</td>
</tr>
<tr>
<td></td>
<td>CheckMate 066</td>
<td>Nivolumab</td>
<td>Dacarbazine</td>
<td>Gamma</td>
</tr>
<tr>
<td></td>
<td>CheckMate 069</td>
<td>Nivolumab + ipilimumab</td>
<td>Ipilimumab</td>
<td>Gamma</td>
</tr>
<tr>
<td></td>
<td>KEYNOTE-006</td>
<td>Pembrolizumab</td>
<td>Dacarbazine or paclitaxel</td>
<td>Log-normal</td>
</tr>
<tr>
<td></td>
<td>METRIC</td>
<td>Trametinib</td>
<td>Dacarbazine</td>
<td>Gamma</td>
</tr>
<tr>
<td></td>
<td>BRIM-3</td>
<td>Nivolumab</td>
<td>Paclitaxel + carboplatin</td>
<td>Gamma</td>
</tr>
<tr>
<td></td>
<td>NSCLC</td>
<td>E4599</td>
<td>Bevacizumab + paclitaxel + carboplatin</td>
<td>Generalized gamma</td>
</tr>
<tr>
<td></td>
<td>PROFILE 1007</td>
<td>Crizotinib</td>
<td>Docetaxel or pemetrexed</td>
<td>Exponential</td>
</tr>
<tr>
<td></td>
<td>TAX 320</td>
<td>Docetaxel</td>
<td>Docetaxel or pemetrexed</td>
<td>Log-logistic</td>
</tr>
<tr>
<td></td>
<td>BR.21</td>
<td>Erlotinib</td>
<td>Docetaxel or pemetrexed</td>
<td>Generalized gamma</td>
</tr>
<tr>
<td></td>
<td>IPASS</td>
<td>Gefitinib</td>
<td>Docetaxel or pemetrexed</td>
<td>Log-logistic</td>
</tr>
<tr>
<td></td>
<td>CA031</td>
<td>Nab-paclitaxel</td>
<td>Carboplatin or paclitaxel</td>
<td>Weibull</td>
</tr>
<tr>
<td></td>
<td>CheckMate 017</td>
<td>Nivolumab</td>
<td>Docetaxel or pemetrexed</td>
<td>Generalized gamma</td>
</tr>
<tr>
<td></td>
<td>CheckMate 057</td>
<td>Nivolumab</td>
<td>Docetaxel or pemetrexed</td>
<td>Generalized gamma</td>
</tr>
<tr>
<td></td>
<td>KEYNOTE-010</td>
<td>Pembrolizumab</td>
<td>Docetaxel or pemetrexed</td>
<td>Generalized gamma</td>
</tr>
</tbody>
</table>

*Continued*
Table S2 (Continued)

<table>
<thead>
<tr>
<th>Trial</th>
<th>Study treatment</th>
<th>Comparator</th>
<th>Parametric curve selected – treatment</th>
<th>Parametric curve selected – comparator</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDA study</td>
<td>Pemetrexed</td>
<td>Cisplatin + gemcitabine</td>
<td>Log-normal</td>
<td>Log-logistic</td>
</tr>
<tr>
<td>PARAMOUNT</td>
<td>Pemetrexed</td>
<td>Placebo</td>
<td>Gamma</td>
<td>Log-logistic</td>
</tr>
<tr>
<td>JMN</td>
<td>Pemetrexed</td>
<td>Placebo</td>
<td>Log-logistic</td>
<td>Log-logistic</td>
</tr>
<tr>
<td>REVEL</td>
<td>Ramucirumab</td>
<td>Docetaxel</td>
<td>Generalized gamma</td>
<td>Gamma</td>
</tr>
<tr>
<td>Renal cell carcinoma</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AXIS</td>
<td>Axitinib</td>
<td>Sorafenib</td>
<td>Weibull</td>
<td>Weibull</td>
</tr>
<tr>
<td>METEOR trial</td>
<td>Cabozantinib</td>
<td>Everolimus</td>
<td>Log-logistic</td>
<td>Log-logistic</td>
</tr>
<tr>
<td>RECORD-I</td>
<td>Everolimus</td>
<td>Placebo</td>
<td>Log-logistic</td>
<td>Log-log normal</td>
</tr>
<tr>
<td>Study 2 FDA label</td>
<td>Lenvatinib + everolimus</td>
<td>Everolimus</td>
<td>Gompertz</td>
<td>Exponential</td>
</tr>
<tr>
<td>CheckMate 025</td>
<td>Nivolumab</td>
<td>Everolimus</td>
<td>Gamma</td>
<td>Gamma</td>
</tr>
<tr>
<td>VEGI05192</td>
<td>Pazopanib</td>
<td>Placebo</td>
<td>Gamma</td>
<td>Gamma</td>
</tr>
<tr>
<td>TARGET</td>
<td>Sorafenib</td>
<td>Placebo</td>
<td>Gamma</td>
<td>Weibull</td>
</tr>
<tr>
<td>A6181034</td>
<td>Sunitinib</td>
<td>IFN-alpha</td>
<td>Gamma</td>
<td>Gamma</td>
</tr>
<tr>
<td>3066K1-304-WW</td>
<td>Temsirolimus</td>
<td>IFN-alpha</td>
<td>Log-normal</td>
<td>Log-normal</td>
</tr>
</tbody>
</table>

Abbreviations: FDA, US Food and Drug Administration; IFN, interferon; NSCLC, non-small cell lung cancer; OS, overall survival; PD-L1, programmed death ligand 1.

Figure S1 Additional persons alive at 1 year vs respective trial comparator (relative to crizotinib) for NSCLC.

Notes: Based on the number needed to treat for an additional person to be alive at 1 year, based on reported cutoffs from each drug’s pivotal trial. Includes FDA-approved interventions in NSCLC for which the relevant Kaplan–Meier overall survival curves were available; trials with comparators of placebo or best supportive care were excluded.

Abbreviations: 1L, first line; 2L, second line; 3L, third line; Bev, bevacizumab; Carbo, carboplatin; Chemo, chemotherapy; Cis, cisplatin; Criz, crizotinib; Doce, docetaxel; Erlot, erlotinib; FDA, Food and Drug Administration; Gefit, gefitinib; Nab-p, nab-paclitaxel; Neci, necitumumab; Nivo, nivolumab; NSQ, nonsquamous; Pac, paclitaxel; Pembro, pembrolizumab; Pemet, pemetrexed; PT, platinum; Ramu, ramucirumab; SQ, squamous.
References


15. Hodi FS, Postow MA, Chesney JA, et al. Clinical response, progression-free survival (PFS), and safety in patients (pts) with advanced melanoma (MEL) receiving nivolumab (NIVO) combined with ipilimumab (IPI) vs IPI monotherapy in CheckMate 069 study. J Clin Oncol. 2015;33:abstr 9004.


73. Hudson TE, Dutuc TC, Ren M, Baig MA, Fishman MN. Subgroup analyses and updated overall survival from the phase II trial of lenvatinib (LEN), everolimus (EVE), and LEN+EVE in metastatic renal cell carcinoma (mRCC). J Clin Oncol. 2016;34:abstr 4553.


140. Douillard JY. The role of EGFR inhibition in mCRC. Presented at: ESMO Preceptorship in Colorectal Cancer 2015; March 28, 2015; Singapore.


