Patient and physician preferences for anticancer drugs for the treatment of metastatic colorectal cancer: a discrete-choice experiment

Juan Marcos González1
Sarika Ogale2
Robert Morlock2
Joshua Posner1
Brett Hauber1
Nicolas Sommer2
Axel Grothey3

1Health Preference Assessment Department, RTI Health Solutions, Research Triangle Park, NC, 2Genentech, South San Francisco, CA, 3Division of Medical Oncology, Mayo Clinic, Rochester, MN, USA

Objective: Many publications describe preferences for colorectal cancer (CRC) screening; however, few studies elicited preferences for anticancer-drug treatment for metastatic CRC (mCRC). This study was designed to elicit preferences and risk tolerance among patients and oncologists in the USA for anticancer drugs to treat mCRC.

Materials and methods: Patients aged 18 years or older with a self-reported diagnosis of mCRC and board-certified (or equivalent) oncologists who had treated patients with mCRC were recruited by two survey research companies from existing online patient panels in the USA. Additional oncologists were recruited from a list of US physicians. Patients and oncologists completed a discrete-choice experiment (DCE) survey. DCEs offer a systematic method of eliciting preferences and quantifying both the relative importance of treatment attributes and the tradeoffs respondents are willing to make among benefits and risks. Treatment attributes in the DCE were progression-free survival (PFS) and risks of severe papulopustular rash, serious hemorrhage, cardiopulmonary arrest, and gastrointestinal perforation. Patients’ and physicians’ maximum levels of acceptable treatment-related risks for two prespecified increases in efficacy were estimated.

Results: A total of 127 patients and 150 oncologists completed the survey. Relative preferences for the treatment attributes in the study were mostly consistent with the expectation that better clinical outcomes were preferred over worse clinical outcomes. Risk tolerance varied between patients and physicians. On average, physicians were willing to tolerate higher risks than patients, although these differences were mostly not statistically significant. Post hoc latent-class analyses revealed that some patients and physicians were unwilling to forgo any efficacy to avoid toxicities, while others were willing to make such tradeoffs.

Conclusion: Differences in preferences between patients and physicians suggest that there is the potential for improvement in patients’ well-being. Initiating or enhancing discussions about patient tolerance for toxicities, such as skin rash and gastrointestinal perforations, may help prescribe treatments that entail more appropriate benefit–risk tradeoffs.

Keywords: metastatic colorectal cancer, discrete-choice experiment, patient preferences, physician preferences, risk tolerance

Introduction

According to the American Cancer Society (ACS), approximately 134,000 people are expected to be diagnosed with colorectal cancer (CRC) and more than 49,000 people are expected to die from CRC in the USA in 2016.1 Two groups of targeted anticancer drugs for metastatic CRC (mCRC), targeting either the vascular endothelial growth factor (VEGF) receptor or the epidermal growth factor receptor (EGFR), have
emerged in the last decade.\textsuperscript{2,3} Currently available agents are the anti-VEGF antibodies bevacizumab, ramucirumab, and afibercept and the anti-EGFR antibodies cetuximab and panitumumab.\textsuperscript{2,3} Results from clinical studies suggest that these agents in combination with standard chemotherapy improve outcomes but are associated with distinct side effects and adverse events.\textsuperscript{2} Thus, physicians and patients must weigh multiple toxicity risks against potentially increased effectiveness when making treatment decisions.

Numerous studies have demonstrated that patients are willing to accept toxicities in exchange for improvements in expected progression-free survival (PFS) in oncology.\textsuperscript{4–6} A few studies have elicited preferences for anticancer drugs for mCRC.\textsuperscript{7–11} No study to date has compared patients’ and physicians’ preferences for anticancer drugs in mCRC using a discrete-choice experiment (DCE).

This study was designed to elicit preferences and risk tolerance for anticancer drugs in mCRC among patients and oncologists in the USA. The study highlights differences in the tradeoffs patients and physicians are willing to make. Understanding these differences may underscore the need for communication between patients and physicians and identify areas in which such communication may be most useful.

**Materials and methods**

**Sample**

Eligible patients were aged 18 years or older with a self-reported diagnosis of mCRC. Physicians were eligible if they were board-certified oncologists (or equivalent) and had treated patients with mCRC. Two survey research companies, AllPoints and Nielsen, recruited patients from their existing online patient panels in the USA between January and May 2014 (Nielsen focused only on the recruitment of patients with mCRC to supplement the study sample). AllPoints also recruited physicians from a list of physicians in the USA treated patients with mCRC. Two survey research companies, AllPoints and Nielsen, recruited patients from their existing online patient panels in the USA between January and May 2014 (Nielsen focused only on the recruitment of patients with mCRC to supplement the study sample). AllPoints also recruited physicians from a list of physicians in the USA during the same time period. Patients and physicians were provided compensation for participating in the study.

The study, including the participating and compensation of patients and physicians, was reviewed by the Office of Research Protection and Ethics at RTI International (the responsible study organization) and approved by its institutional review board. All respondents provided online informed consent.

**DCE survey instruments**

DCEs offer a systematic method of eliciting preferences and quantifying both the relative importance of treatment attributes in the choice of treatments and the tradeoffs patients and physicians are willing to make between the benefits and risks of treatments.\textsuperscript{12} DCEs have been used to elicit patient and physician treatment preferences separately in numerous therapeutic areas.\textsuperscript{13–16} DCEs are based on the principle that treatments are bundles of attributes or features (for example, efficacy and toxicities) and the value of a given treatment can be decomposed into the relative value ascribed to each treatment attribute.\textsuperscript{17} Choices between treatments reflect the relative value of the treatment attributes,\textsuperscript{18} and systematic analysis of these choices under appropriate experimental control can produce attribute-specific weights indicating the importance of each attribute in respondents’ choices.\textsuperscript{17,19}

In this study, two DCE surveys were developed and administered online: one for patients and one for physicians (Supplementary material). The surveys were developed following good research practices.\textsuperscript{20} The DCE included a series of choice questions, each asking the respondents to choose between two hypothetical mCRC treatments defined by the attributes presented in Table 1. These attributes were selected based on prescribing information for anti-VEGF and anti-EGFR antibodies and individual consultation with five oncologists. The five oncologists were interviewed via telephone and asked to evaluate 19 outcomes associated with

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Levels</th>
<th>Clinically relevant range</th>
</tr>
</thead>
<tbody>
<tr>
<td>PFS (how long before the cancer grows)</td>
<td>12 months, 8 months, 6 months</td>
<td>6–10 months</td>
</tr>
<tr>
<td>Chance of severe papulopustular rash (chance of severe skin rash)</td>
<td>None, 5% (5 out of 100), 10% (10 out of 100), 25% (25 out of 100)</td>
<td>0–20%</td>
</tr>
<tr>
<td>Chance of serious hemorrhage (chance of serious bleeding)</td>
<td>None, 2% (2 out of 100), 10% (10 out of 100), 35% (35 out of 100)</td>
<td>0–5%</td>
</tr>
<tr>
<td>Chance of cardiopulmonary arrest (chance of severe heart attack)</td>
<td>None, 2% (2 out of 100), 10% (10 out of 100), 20% (20 out of 100)</td>
<td>0–2%</td>
</tr>
<tr>
<td>Chance of GI perforations</td>
<td>None, 2% (2 out of 100), 10% (10 out of 100), 20% (20 out of 100)</td>
<td>0–2%</td>
</tr>
</tbody>
</table>

**Table 1 Attributes and levels**

**Note:** Attribute terms in parentheses correspond to the attribute name in the patient survey. **Abbreviations:** GI, gastrointestinal; PFS, progression-free survival.
anti-VEGF and anti-EGFR antibodies. The oncologists were then asked to select and rank the most important outcomes when making prescribing decision for mCRC patients among the outcomes presented. The most highly rated outcomes from this exercise were included in the survey instrument. All attributes were described to respondents prior to the choice questions in the DCE. The descriptions of all attributes are presented in the Supplementary material.

The salience and clarity of the attributes and attribute levels were assessed through in-person semistructured pretest interviews with patients and physicians. Individuals were invited to participate in one of 15 patient pretest interviews if they reported having a physician’s diagnosis of CRC, with five having a self-reported diagnosis of mCRC. Physicians were invited to participate in one of five pretest interviews if they were board-certified practicing oncologists or gastroenterologists who treat patients with mCRC. In each interview, respondents were asked to follow a think-aloud protocol as they completed the instrument. Respondents were also asked to complete a series of mock choice questions intended to assess whether they understood and were willing to make tradeoffs between hypothetical mCRC treatments described using the study attributes. The levels for the attributes were selected to encompass the clinically relevant manifestations of the attributes in currently available mCRC treatments. Based on respondents’ answers to the mock choice questions during the pretest interviews, the attribute levels were revised to cover scenarios that may not be currently technically feasible but would represent meaningful tradeoffs to respondents.

In the final survey instrument, choice questions included pairs of hypothetical treatment profiles. The hypothetical profiles and profile pairs in the choice questions were prepared following an experimental design that systematically varied the levels for each attribute in Table 1. The systematic variation in attribute levels permits the use of respondents’ choices to infer which attribute differences across treatments drive respondents’ treatment preferences. Good research practices were used to construct a fractional factorial experimental design with an SAS implementation of a commonly used D-optimal algorithm. Random-parameters logit (RPL) was used to estimate relative preferences for all attribute levels in the study based on respondents’ answers to the series of choice questions. Random-parameters logit (RPL) relates the probability of choosing one treatment over another to observable differences between treatment options. The model also takes into account the distribution of preferences among respondents in the sample (i.e., preference heterogeneity). Parameter estimates from an RPL model can be interpreted as preference weights, which represent the relative contribution of changes in treatment attributes to choice. A higher preference weight for a specific attribute level indicates a greater likelihood that a treatment including that attribute level will be chosen, all else being equal.

The average change in respondents’ preference for a treatment induced by changes in attribute levels represents the relative importance of the attribute change. Because all preference weights are estimated on the same scale, it is possible to directly compare the relative importance of a change in the levels of one attribute and the relative importance of a change in the levels of another attribute. The ratio of the importance of these changes provides a measure of the rate at which respondents are willing to tradeoff among attributes.

Attribute levels were specified as effects coded, categorical variables. Patient and physician data were analyzed separately. The patient model controlled for potential differences in the model variance between respondents who were asked to participate through different recruitment strategies.

Maximum acceptable risk

Maximum acceptable risk (MAR) is the increase in the likelihood of an adverse event that yields a relative importance equal to the relative importance of a prespecified increase in efficacy. It is a measure of the maximum level of risk that
respondents would be willing to accept to achieve a specific increase in treatment benefit.\textsuperscript{12,25} For both patients and physicians, we calculated the MAR for each toxicity, given two different levels of improvement in PFS – an increase in PFS from 6 months to 8 months and from 8 months to 12 months.

Post hoc analysis
Although RPL models account for heterogeneity in respondents’ choices and estimated respondents’ preferences, RPL does explain systematic variations in preferences among subgroups or segments in the sample. To evaluate such systematic variations in preferences more closely, we conducted a number of post hoc analyses. Differences in the incidence of CRC and treatment decisions based on patients’ age and gender have been documented in the past.\textsuperscript{26–29} Therefore, we conducted subgroup analysis to determine whether age and gender were associated with systematic differences in patient preferences. We also conducted latent-class analyses to determine whether segments with systematically different preferences existed within the patient and physician samples.\textsuperscript{30}

Results
Sample
A total of 302 patients who were expected to meet the study inclusion criteria responded to the invitation to participate in the study. Of those patients who responded, 157 (52\%) were eligible and consented to participate. Of those who were eligible and consented to participate, 127 (81\%) completed the survey. A total of 211 physicians who were expected to meet the study inclusion criteria responded to the invitation to participate in the study. Of the physicians who responded, 183 (87\%) were eligible and consented to participate. Of those who were eligible and consented to participate, 150 (82\%) completed the survey.
Table 2 presents patients’ demographic and disease characteristics and physicians’ demographic characteristics and experience treating mCRC. Mean (median) patient age was 45.5 (48) years. Most patients were male (54%) and diagnosed at least 1 year ago (59%). The average physician in the sample was aged 47 years. Approximately one-third of physicians had been in practice for more than 15 years (38%) and treated between 21 and 30 patients per month (33%).

### Preference weights
Figures 2 and 3 present the estimated preference weights (95% confidence interval) for all patients and all physicians, respectively. For patients, across most attribute levels, preference weights changed in the way expected; that is, better clinical outcomes were preferred to worse clinical outcomes. The only exception was for severe papulopustular rash, where the mean preference weight estimates suggest that a 10% chance of severe papulopustular rash would be preferred over a 5%
chance. However, the disordered preference weights were not statistically significantly different from each other, suggesting that the disordering reflects a lack of precision in those specific estimates. Among physicians, better clinical outcomes were also preferred to worse clinical outcomes. Exceptions to this pattern were observed for severe papulopustular rash and gastrointestinal perforation. As with patients, disordered preference weights were not statistically significantly different from each other.

Among all patients, an improvement in PFS from 8 months to 12 months had a relative importance of 0.80 (0.73 to 0.87). Reducing the risk of treatment-related cardiopulmonary arrest from 2% to 0% had a relative importance of 0.23 (1.20–0.97). Therefore, increasing PFS from 8 months to 12 months was approximately 3.5 times as important (0.80/0.23) as reducing the treatment-related risk of cardiopulmonary arrest from 2% to 0%. An improvement in PFS from 6 months to 8 months (0.61 = −0.06 to −0.67) was approximately 2.7 times as important to patients as eliminating a 2% treatment-related risk of cardiopulmonary arrest.

Also among physicians, an improvement in PFS from 8 months to 12 months had a relative importance of 1.5 (1.4

---

### Table 2 (Continued)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Statistic or category</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Which of the following describes your practice? (check all that apply)</td>
<td>Office-based private practice</td>
<td>97 (64.7)</td>
</tr>
<tr>
<td>Hospital-based private practice</td>
<td>33 (22.0)</td>
<td></td>
</tr>
<tr>
<td>Academic hospital-based practice</td>
<td>30 (20.0)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>5 (3.3)</td>
<td></td>
</tr>
<tr>
<td>Approximately how many patients with metastatic colorectal cancer do you treat monthly?</td>
<td>5 patients or fewer</td>
<td>11 (7.3)</td>
</tr>
<tr>
<td>6–10 patients</td>
<td>20 (13.3)</td>
<td></td>
</tr>
<tr>
<td>11–20 patients</td>
<td>36 (24.0)</td>
<td></td>
</tr>
<tr>
<td>21–30 patients</td>
<td>50 (33.3)</td>
<td></td>
</tr>
<tr>
<td>More than 30 patients</td>
<td>33 (22.0)</td>
<td></td>
</tr>
<tr>
<td>Which of the following targeted treatments have you used to treat patients with mCRC? (check all that apply)</td>
<td>Bevacizumab</td>
<td>147 (98.0)</td>
</tr>
<tr>
<td>Cetuximab</td>
<td>138 (92.0)</td>
<td></td>
</tr>
<tr>
<td>Panitumumab</td>
<td>122 (81.3)</td>
<td></td>
</tr>
<tr>
<td>Ziv-aflibercept</td>
<td>88 (58.7)</td>
<td></td>
</tr>
<tr>
<td>Regorafenib</td>
<td>108 (72.0)</td>
<td></td>
</tr>
<tr>
<td>Other targeted agent</td>
<td>16 (10.7)</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>1 (0.7)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: *One respondent did not answer this question. *Nine respondents did not answer this question.

Abbreviations: mCRC, metastatic colorectal cancer; SD, standard deviation.

---

**Figure 2** Patient preference weights (N = 127).

Abbreviations: GI, gastrointestinal; PFS, progression-free survival.
Reducing the risk of treatment-related cardiopulmonary arrest from 2% to 0% had a relative importance of 0.7 (1.9–1.2). Therefore, increasing PFS from 8 months to 12 months was approximately 2.5 times as important (1.5/0.7) as reducing the risk of treatment-related cardiopulmonary arrest from 2% to 0%. An improvement in PFS from 6 months to 8 months (1.2 = −0.1 to [−1.3]) was approximately 1.7 times as important to physicians as eliminating a 2% treatment-related risk of cardiopulmonary arrest.

Maximum acceptable risk

Table 3 presents MAR estimates for patients and physicians for two improvements in efficacy for each of the four risks shown in the study.

On average, physicians were more tolerant than patients of each of the four risks in the study. For example, physicians would accept up to a 16% chance of serious hemorrhage in exchange for a 2-month improvement in PFS (from a baseline of 6 months), while patients would only accept up to an 8% chance of serious hemorrhage for the same change in PFS. While mean differences in levels of MAR were systematically higher for physicians than patients, these differences were not statistically significant at the 95% confidence level.

Post hoc analysis

We found no statistically significant differences in patient preferences based on gender (p = 0.10). However, respondents above and below the median sample age of 48 years had statistically significantly different preferences (p = 0.01). Older patients placed greater weight on improving PFS and avoiding gastrointestinal perforation and skin rash than did younger patients, while the relative importance of avoiding a heart attack and serious bleeding was the same in both subgroups.

Latent-class analyses revealed two distinct classes of patients and two distinct classes of physicians (data not shown). Each sample included one class that apparently focused almost entirely on PFS improvements and thus was unwilling to forgo any PFS to avoid toxicities. Although respondents who are likely to be in this class provided primarily information about their preference for PFS improvements,
their responses also suggest indifference between the toxicities in the study. The average probability of being in this class was approximately 28% in the patient sample. Among patients who appeared more willing to tradeoff between PFS and toxicities, risk tolerance estimates for any toxicity given increases in PFS were lower than the corresponding estimates of risk tolerance in the full sample. Respondents with no history of a heart attack and those who had greater difficulty with the choice questions were more likely in the class focusing almost entirely on PFS.

The average probability of being in the class in which respondents appeared to be unwilling to forgo any PFS to avoid toxicities was 33% in the physician sample. Similarly to patients who were more willing to tradeoff between PFS and toxicities, physicians in this group were less risk tolerant than physicians in the full sample. Class membership in the physician sample was not explained in a statistically significant way by known physician characteristics.

As in the full sample, large differences in the mean level of risk tolerance existed between patients and physicians among respondents who are willing to trade between PFS and toxicities.

**Discussion**

Many studies have evaluated patients’ preferences for the management of mCRC. However, only two previous studies have elicited patients’ or physicians’ preferences for anticancer drug treatments for mCRC using a DCE, and our study is the first to contrast patient and physician preferences for attributes of anticancer drugs in mCRC using this methodology. Relative preferences for the treatment features in the study were mostly consistent with the expectation that better clinical outcomes were preferred over worse clinical outcomes. In cases where mean preferences did not follow this pattern, differences were not statistically significant, suggesting indifference between the disordered attribute levels. This result highlights that relative preferences did not change linearly with outcome severity or likelihood, suggesting that the value of treatment improvements changed with the baseline attribute levels. In turn, this implies that the impact of mCRC therapies on patients may depend not only on the clinical changes induced by treatment but also on the patient’s current health status.

As expected, risk tolerance depended on the benefits considered: longer delays in tumor growth were consistently associated with higher MARs. However, risk tolerance varied between patients and physicians. For patients, average MAR values suggest that the relative importance of side effects can change with the magnitude of benefit. The change in average MAR across risks is the result of nonlinear growth in attribute importance with changes in the probability of experiencing an adverse event.

Our results indicate that, for all treatment-related risks, physicians were willing to tolerate greater probability of adverse events than patients. Previous comparisons of patients’ and physicians’ risk tolerance in other disease areas have shown greater risk tolerance among physicians when efficacy improvements alleviate very severe disease symptoms, but not so when the improvements deal with milder symptoms. Our results are consistent with these findings. The results also suggest that treatment benefits among patients with mCRC may be enhanced by better aligning patients’ and physicians’ views about the efficacy and risk of adverse events associated with mCRC treatments.

While, on average, patients and physicians were willing to forgo some level of increase in PFS to avoid treatment-related toxicities, post hoc analysis of the data collected in this study indicated that this was not necessarily true for all patients or all physicians. For a minority of patients and a minority of physicians, risk tolerance was too large to measure, as these respondents would appear to be willing to accept any of the risk levels included in the DCE to prolong PFS. Thus, the mean estimates of MAR for the full samples of patients and physicians may underestimate the level of risk tolerance for some patients and physicians and overstate the level of risk tolerance for others. Nevertheless, without evidence suggesting that risk tolerance beyond the risk levels in the study is a biased representation of respondents’ preferences, the average MAR results ought to include all views elicited in the DCE survey.

This DCE survey asked patients and physicians to evaluate hypothetical mCRC treatments. Potential hypothetical bias was minimized by the survey offering alternatives that mimic real-world tradeoffs as closely as possible. However, not all attributes of treatments were included in the survey, and differences can arise between stated and actual choices. Similarly, the results from this study relate only to the attributes and attribute levels presented in the survey. Alternative definitions of the attributes and attribute levels could also affect the results.

Also, the study design did not allow us to identify variations in physicians’ preferences with patient characteristics. The hypothetical patient for whom physicians were asked to make decisions in the survey differed from the average respondent in the patient sample in observed ways, such as patient age. These differences may have influenced the discrepancy in risk tolerance observed between patients and physicians. The preference results for patients by age groups, however,
suggest that risk tolerance for gastrointestinal perforations and papulopustular rash among patients who are closer in age to the hypothetical patient evaluated by physicians is not closer to physicians risk tolerance. Older patients were more concerned about these two toxicities, implying that patients' risk tolerance for these toxicities decreases as their age increases. Future work should evaluate this result more formally by designing hypothetical patient profiles that differ in age and assess whether this specific patient characteristic matters to physicians recommending mCRC treatments for their patients.

Another limitation is that the sampling process was not designed or weighted to ensure representativeness of the population of patients with mCRC or the physicians treating patients with mCRC. Thus, the results offered here should be evaluated cautiously as it is not possible to determine to what extent they provide a representative view of patients with mCRC or the physicians who treat them. Additional work eliciting patient and physician preferences should be undertaken and reported in the future.

The sizes of the patient and physician samples in this study are small relative to other DCE studies in the literature,15 which may have some implications for the interpretation of our results. We cannot know whether the lack of significance in differences between preference weights represents indiffERENCE or is a result of small sample sizes. Nevertheless, we were able to evaluate post hoc the ability of the experimental design to identify these differences given the actual sample sizes, based on the statistical properties of the experimental design and the estimated differences in preference weights.33 This leads us to conclude that the lack of statistical significance in the differences between preference weights for any remaining attributes is not an artifact of the sample sizes.

In addition, we cannot know whether the lack of statistical significance in differences between MARs of patients and physicians represents a true lack of difference or is a result of our sample sizes. However, given the large differences between estimated patient and physician MARs, we surmise that the lack of statistical difference is most likely due to sample size. Finally, our sample sizes are likely insufficient to estimate precise results using latent-class analysis, especially for the classes of patients and physicians who represent a minority of the sample and appear to be unwilling to trade among attributes in this survey. However, the results of this analysis provide a clear indication of differences in preferences between patients and physicians and within the patient and physician samples. Future research in this area should focus on quantifying and understanding such differences.

The results of this study suggest that it is important for physicians to discuss their views on the tradeoffs between the benefits and risks of mCRC treatments with their patients, in addition to understanding patients’ views of these tradeoffs. Such discussions may help patients appreciate the balance between achieving the best possible tumor response and managing toxicities and may enable physicians to understand patients’ goals and priorities and thus determine a treatment approach that is right for each patient.

**Acknowledgments**

This study was sponsored by Genentech. The authors acknowledge Amy Pugh for assisting with the analysis in this study.

**Disclosure**

Brett Hauber and Joshua Posner are employees of RTI Health Solutions, which received research funding from Genentech. Juan Marcos González was an employee of RTI Health Solutions at the time of the study. Nicolas Sommer and Sarika Ogale are employees of Genentech; Robert Morlock was an employee of Genentech at the time of the study. The authors report no other conflicts of interest in this work.

**References**


