The European Journal of Health Economics

Quality of Life Utility-Core 10 Dimensions (QLU-C10D) scores based on cancer patients' health preferences are more sensitive than those based on general public's health preferences

--Manuscript Draft--

Manuscript Number:	EJHE-D-25-00109R1
Full Title:	Quality of Life Utility-Core 10 Dimensions (QLU-C10D) scores based on cancer patients' health preferences are more sensitive than those based on general public's health preferences
Article Type:	Original Research
Keywords:	Value Set, Cancer Patients, General Population, QLU-C10D, China
Corresponding Author:	Weidong Huang Harbin Medical University CHINA
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Harbin Medical University
Corresponding Author's Secondary Institution:	
First Author:	Yiyin Cao
First Author Secondary Information:	
Order of Authors:	Yiyin Cao
	Haofei Li
	Xu Jin
	Madeleine T. King
	Richard Norman
	Ling Jie Cheng
	Hongjuan Yu
	Weidong Huang
	Nan Luo
Order of Authors Secondary Information:	
Funding Information:	National Natural Science Foundation of China (72274045) — Professor Weidong Huang
	National Natural Science Foundation of China (71974048) — Professor Weidong Huang
Abstract:	Objective: This study aimed to develop a value set for the European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Utility-Core 10 Dimensions (QLU-C10D) based on the preferences of Chinese cancer patients, and to compare results with the Chinese general population value set. Methods: Patients with mixed cancer diagnoses were recruited from two hospitals in China between September and November 2024. Patients completed the EORTC Core Quality of Life Questionnaire (QLQ-C30) followed by a discrete choice experiment (DCE) valuation survey consisting of 16 choice sets through face-to-face interviews. A conditional logit model was applied to analyze patient preferences. The characteristics and sensitivity of the derived QLU-C10D value set were compared with those of the Chinese general population QLU-C10D value set. Sensitivity was evaluated by ability

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

	to discriminate cancer stage and performance status, and effect sizes (ES) were estimated. Results: Data from 659 cancer patients were analyzed. The cancer patients' value set showed a broader utility range (-0.092 to 1.0) compared to the general population's value set (0.083 to 1). The mean utility score for the cancer patient sample (0.656) was significantly lower than that for the general population (0.738, $P < 0.01$). Both value sets demonstrated good sensitivity, but the cancer patients' value set was more effective in distinguishing clinical subgroup differences (ES = 2.76–4.46 vs. ES = 2.47–3.48). Conclusion: Differences in value sets between cancer patients and the general population underscore the importance of more comprehensively integrating patient perspectives into existing evaluation frameworks.
Response to Reviewers:	Dear Editor, We sincerely appreciate the editor's kind attention and voluntary work in handling our manuscript. Thank you for your time and support. Reviewers' comments: Reviewer #1: Dear Reviewer ： Thank you from the bottom of our hearts for your extraordinarily thorough and insightful review of our manuscript. Your attention to even the most detailed aspects—including the significance of statistical coefficients and the clarity of figures—reflects a level of rigor that we deeply admire and appreciate. We are particularly grateful for your recognition of our efforts in recruiting a large number of cancer patients within a short timeframe; this was indeed a challenging feat, as we had only a limited window of approval from the participating hospitals. To meet this timeline, our team mobilized significant human resources—including trained interviewers, clinical collaborators, and data managers—working closely to streamline recruitment, ensure participant comfort, and maintain data quality. Your acknowledgment of this effort means a great deal to us. We wholeheartedly agree with your observations: there were indeed points in our manuscript that lacked sufficient clarification, and revisions to the figures are absolutely necessary to enhance readability and precision. Your feedback has pinpointed exactly where we need to strengthen the work, and we are committed to addressing every single one of your suggestions with the utmost care. Rest assured, we will revise the manuscript diligently based on your expert guidance—clarifying ambiguous details, refining the figures to better convey key findings, and ensuring all analyses are presented with the rigor they deserve. We believe this process will significantly elevate the quality of our work, and we owe this to your thoughtful input. Thank you again for your time, expertise, and dedication to helping us improve. Below is our point-by-point response, with revised sections highlighted in red for easy reference. 1. The authors need to be very careful with using the terminology on the concepts presented correctly and coherently. I suggest using the term "value sets" for anything that refers to the utility decrements of the patients or general populations. Be careful to not confuse the terms QALYs, utility values, value sets, utility decrements, and preferences. The field needs to be very careful in using these terms correctly as these concepts are all similar but have a critical distinction. Reply：Thank you profoundly for this critical and insightful feedback. We fully agree that precise and consistent use of terminology is foundational to clarity in health economics and preference research, and we deeply appreciate you highlighting this important point—your guidance will significantly enhance the rigor of our manuscript. We recognize that our initial use of terms such as “value sets,” “QALYs,” “utility values,” “utility decrements,” and “preferences” may not have consistently reflected

	their nuanced distinctions, and we take this concern very seriously. To address this, we have conducted a systematic review of the entire manuscript (including text, figures, tables, and references) to ensure each term is used with strict adherence to its established definition in the field: “Value sets” will specifically refer to the preference-based scoring algorithms derived from the respective populations (cancer patients and the general population) used to quantify health states. “Utility values” will denote the single-index scores assigned to health states (e.g., from 0 to 1) using these value sets. “Utility decrements” will describe the differences between these values (e.g., from a reference state to an impaired state), capturing the magnitude of health-related quality of life loss. “QALYs” will be reserved for the metric integrating utility values over time, reflecting the overall health benefit of an intervention. “Preferences” will refer to the underlying judgments (e.g., trade-offs) guiding these valuations, as elicited through methods like discrete choice experiments. Notably, we have revised the titles of key figures to align with these definitions, moving away from previous references to “utility decrements.” Specifically: Fig. 1 is now titled “QLU-C10D value sets for the cancer patient sample and the general population,” directly focusing on the scoring algorithms themselves. Fig. 2 has been updated to “Difference in value sets between cancer patients and the general population,” explicitly emphasizing the comparison of these algorithmic frameworks. These adjustments extend to the entire text, where descriptions of analyses and results are strictly aligned with the defined terms. All revisions—whether in figures, text, or tables—are marked in red for easy reference. Thank you again for prompting us to refine these critical distinctions—your expertise has been invaluable in elevating the quality of our work. 2. One of the main conclusions the authors wish to draw is that patients' value sets lead to lower utility values compared to the value sets of the general population (this is often stated). While I understand that this claim is supported by applying the value sets on hypothetical health states of the QLU-C10D descriptive system, it lacks to inform the reader how these hypothetical health states were defined. As the authors present this as the main finding, detailed information on the definition of the hypothetical health states is required to inform the readership. Reply：Thank you sincerely for highlighting this critical point. We fully agree that detailed information on the definition of hypothetical health states is essential to contextualizing our main conclusion, and we appreciate your guidance in strengthening the transparency of our analysis. As suggested, we have supplemented the third paragraph of Section 2.5 (Data Analysis) with specific details on the hypothetical health states (revisions marked in red): " Additionally, 1,000 hypothetical health states were generated by randomly combining the 4 levels across the 10 dimensions of the QLU-C10D descriptive system. This sampling approach ensured proportional representation of each dimension’s levels, with states spanning the full severity spectrum—from mild (e.g., most dimensions at levels 1–2) to severe (e.g., majority at levels 3–4)—and explicitly including the best-possible (all dimensions at level 1) and worst-possible (all dimensions at level 4) health states. This comprehensive coverage of potential health states was designed to ensure robust utility comparisons between the two value sets." We believe these details adequately clarify how the hypothetical health states were defined and justify the generalizability of our findings. Thank you again for prompting this important improvement—your input has significantly enhanced the rigor of our methodology.

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

	3.The general population value sets for mainland China were weighted for education and "geography" (I believe to achieve representation of persons in different provinces). This study does not weight the results. I understand that it is not possible to do so for "geography" but would be possible for the variable "education". I would recommend doing so (to reflect a similar distribution as in the gen. pop. publication), if the authors wish to abstain from that I would suggest adding this in the limitation section. Reply：Thank you sincerely for this thoughtful suggestion regarding education weighting—your insight has helped us reflect more deeply on the alignment between our methodology and the general population value set, and we truly appreciate you pushing us to clarify this important methodological choice. We completely agree that education weighting is a valuable practice for general population studies, and we took your recommendation very seriously: in fact, we explicitly tested this approach during our pre-analysis phase (as you'll see in our revised limitations section) to ensure we weren't overlooking a critical step. As we worked through the analysis, we faced a few key considerations that led us to prioritize unweighted results, and we'd like to share them with you in the spirit of transparency: First, the demographic profile of cancer patients in China—particularly age and education distributions—differs inherently from the general population (e.g., with higher representation in older age groups, who often have distinct educational backgrounds). Unfortunately, there's a lack of robust, disease-specific demographic benchmarks for China's cancer patient population, making it challenging to derive weighting criteria that truly reflect the realities of this group. Second, our sampling strategy focused on high-incidence cancer types to ensure our findings would resonate with the most common clinical scenarios in China's oncology practice. In preliminary tests, we found that forcing alignment with general population education distributions risked diluting the unique preferences of these patients—especially differences in how they value health states across the cancer types we prioritized. Since our core goal was to capture patient-specific value sets, we worried this might undermine the ecological validity of our work. We wholeheartedly agree that this is a limitation, and we've added a detailed note to the limitations section (marked in red) to acknowledge it: "Thirdly, we did not apply education weighting (as used in the general population value set) due to inherent differences in education distributions between cancer patients and the general population; this omission may weaken the stability of extrapolating our findings for cross-group comparisons." Importantly, while we opted not to apply education weighting in this manuscript, we fully recognize the significance of educational background in shaping health preferences—and this is precisely why we've dedicated a companion study to exploring such nuances in depth. Developed in parallel with the current work, this companion manuscript focuses on unpacking preference heterogeneity among cancer patients, with a specific emphasis on how educational attainment intersects with other factors (e.g., cancer type, age, and treatment modality) to influence health state valuations. Tentatively titled "Exploring Preference Heterogeneity in QLU-C10D Health State Valuations: Variations Across Cancer Types and Treatment Modalities," this study uses mixed logit modeling and interaction terms to dissect subgroup differences. For instance, preliminary analyses reveal that within the same cancer type, patients with different educational backgrounds may prioritize certain health attributes differently—such as greater emphasis on "social function" among those with higher education, compared to stronger focus on "physical comfort" among those with lower education. These insights build on the findings of the current manuscript, offering a more granular understanding of how education, alongside other variables, shapes patient preferences. This companion study aligns closely with your suggestion to explore the role of education, and we believe it provides the depth of analysis this topic deserves. We plan to reference its key findings in the discussion section of the current manuscript to

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

contextualize our unweighted results and highlight avenues for future research.

Thank you again for prompting this careful consideration—your attention to this detail has strengthened the rigor of our methodology, and we're grateful for the opportunity to clarify our approach. We hope this explanation addresses your concern, and we'd be delighted to share more details about the companion study if you're interested. Please feel free to reach out for further discussion!

4.The manuscript would benefit from a proofreading service. This would help to align the consistency of the statements made and help with the use of the terminology as well.
Reply：Thank you for your suggestion regarding proofreading. We appreciate your attention to the consistency of statements and terminology.

Notably, several authors on our team are native English speakers, and they have further revised and polished the manuscript on the basis of previous modifications to ensure linguistic accuracy, terminological consistency, and overall coherence.

We believe these efforts have enhanced the quality of the manuscript, and we thank you again for your valuable feedback.

5.The interpretation of the results should be further done taking measures of statistical uncertainty into account. The standard errors presented in the tables sometimes exceed the difference of the utility decrements of the gen. population and the patients. This is not taken into account when presenting the results.
Reply：We sincerely appreciate your insightful feedback on incorporating statistical uncertainty into result interpretation.

To address your concern, we have revised our analysis and visualization to explicitly account for statistical uncertainty:

We calculated the standard errors for cancer patient utility decrement and integrated them into Figure 1 as error bars. These error bars directly reflect statistical variability, and we have updated the textual description of the figure to clarify this:“ The standard errors for cancer patients' utility decrement—visualized as error bars in Figure 1—occasionally exceed these observed differences, with such error bars reflecting statistical uncertainty.”

By visualizing standard errors, we ensure that readers can now interpret utility decrement differences (e.g., cases where standard errors surpass these differences) with full awareness of statistical uncertainty. This addition helps contextualize the reliability of our findings and aligns with your recommendation to consider precision in result interpretation.

Abstract:
6. Page 1, line 29-30: "For China, utility scores based on the general population value set underestimate quality adjusted life years (QALY) compared to those based on the cancer patient value set." This conclusion is not supported by the analysis. 1.) you did not calculate QALYs, 2.) utility values are higher in the gen. pop. Potentially leading to an overestimation of QALYs.
Reply：Thank you for your astute observation regarding our inappropriate conclusion about QALYs. We fully agree with your points: we did not calculate QALYs in this study, and the higher utility values in the general population value set could potentially lead to overestimation of QALYs, making our original claim unsupported.

To address this, we have revised the problematic statement to strictly align with our analysis, which focuses on utility value sets rather than QALY calculations. The updated conclusion now reads:

“Differences in value sets between cancer patients and the general population underscore the importance of more comprehensively integrating patient perspectives into existing evaluation frameworks.”

This revision removes all references to QALYs, anchors the conclusion in our core

findings (comparisons of utility value set properties), and emphasizes the practical relevance to cost-effectiveness analyses—consistent with the scope and results of our study.

Thank you again for guiding us to refine this conclusion, enhancing the accuracy and rigor of our work.

Introduction:
7. Page 2, line 8: Consider revising "HRQoL utility values" to "Country specific value sets, reflecting societal preferences"
Reply：Thank you for your suggestion to revise "HRQoL utility values". We have followed your advice and revised the phrase to "Country specific value sets, reflecting societal preferences" as recommended. This change aligns with your insight and improves the precision of the statement. Thank you again for your valuable input.

8.Page 2, line 17: consider revising "patient preferences" to "value sets represent the preferences informed by the lived experience".
Reply：We have adopted it and revised the content to: "Therefore, integrating value sets representing preferences informed by the lived experience alongside societal perspectives offers a more comprehensive framework for clinical and policy decision-making." Your guidance helps improve the precision of our expression.

9.Page 2, line 19: I do not believe that Germany relies on QALY assessment at all. Read IQWIG guidelines for a proper understanding of the German HTA strategies.
Reply：Following your advice, we have revised the relevant text to: "In Germany, the HTA framework led by the Institute for Quality and Efficiency in Health Care (IQWiG) focuses on qualitative assessments of 'added clinical benefit' from the patient perspective, such as survival gains and symptom improvement. It directly incorporates patient-reported outcomes (PROs) to evaluate therapies against comparators, ensuring patient voices are central to treatment benefit evaluations."

This revision removes any implication of QALY reliance and aligns with the actual focus of Germany's HTA strategy, as highlighted in your comment. Thank you again for helping us improve the accuracy of our work.

10.Further I want to provide an excerpt from the Swedish value set publication:
"Sweden relies on cost-effectiveness analysis in health economic appraisal, but this analysis is subordinate to the other two guiding principles of human value and need and solidarity when making resource-allocation decisions [1, 4]. In general, the use of CUA and QALYs is recommended by the Dental and Pharmaceutical Benefits Agency [41]."
Reply：Thank you for providing the excerpt from the Swedish value set publication, which offers valuable context for Sweden’s health economic appraisal approach. We have incorporated this insight while retaining our focus on patient-centered elements in the revised text:

" Similarly, Sweden’s Dental and Pharmaceutical Benefits Agency (TLV) emphasizes ethical principles like "human value and need" and "solidarity" in its health economic evaluations, which prioritize patient experiences. These evaluations integrate condition-specific PROs to align resource allocation with patients’ actual needs [8]."

This revision stays true to the core of our focus on patient preferences and the role of ethical principles in TLV’s framework, as highlighted in your reference. We appreciate your guidance in ensuring the accuracy of this section.

Additional Minor Points:
11. Page 2, line 37-38: English: "of universally applying…."
Reply：We have revised the text to "of universally applying…" as suggested, which aligns better with standard English usage. Your careful review is greatly appreciated, as it contributes significantly to enhancing the precision and professionalism of our manuscript.

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

12.Page 2, line 44-45: Correct full name: Quality of Life Questionnaire - Core 30
Reply：Thank you for correcting the name "Quality of Life Questionnaire - Core 30" on Page 2, lines 44-45. We apologize for this oversight in our field and have revised it as advised. Your careful review is much appreciated.

13.Page 3, line 1: Here and in the discussion you write: "with the highest global cancer incidence". If this is true, please provide a reference to the statement (e.g. GLOBOCAN, IARC data).
Reply：As advised, we have supplemented the relevant parts with a specific reference to validate the statement: Han B, Zheng R, Zeng H, Wang S, Sun K, Chen R, Li L, Wei W, He J: Cancer incidence and mortality in China, 2022. J Natl Cancer Cent 2024, 4(1):47-53. Specifically, this reference has been added and highlighted in red both in the corresponding part of the introduction and in the first paragraph of the discussion where the same statement appears.

Valuation Survey and DCE Design:
14. Page 4, line 7: you write "five of ten HRQoL dimensions" - typically four HRQoL dimensions + survival time. Also the dimension "physical functioning" is presented as two items in the conventional DCE tasks (displaying 10 dimensions, but 11 items). Please correct for accurate writing.
Reply：Thank you for your careful observation regarding the description of HRQoL dimensions and attributes in the DCE tasks. We appreciate you pointing out the need to clarify that the conventional design typically includes four HRQoL dimensions plus survival time, as well as the distinction between the 10 domains (with "physical functioning" split into 2 attributes, totaling 11 items).

We have revised the text to accurately reflect these details:

" To control biases, Health State A/B assignment and the order of 11 attributes across 10 QLU-C10D domains (the physical functioning dimension was split into 2 attributes, one for ‘long walk’ and one for ‘short walk’ (see Table A) were randomized per respondent; a balanced incomplete block design (BIBD) defined four differing dimensions per set (others held equal) to minimize burden. Each set presented two hypothetical QLU-C10D health states, differing in four of ten HRQoL dimensions and survival times (1, 2, 5, or 10 years), with differences highlighted for clarity. Examples are shown in Figures A and B (Chinese version) in the supplementary material. Details on QLU-C10D methodology, including formats, attribute ordering, and reliability, are provided in related studies [10, 57-59]."

This revision aims to align with the conventional structure of DCE tasks and accurately clarify the relationship between domains, attributes, and dimensions. Thank you again for helping us refine the precision of this description.

Results:
15. Table 2: You fail to indicate which coefficients are statistically significant. This is indicated in the footnote of the table, not a single coefficient is marked with an .
Reply：We have added the relevant markers as advised, with the revised sections in Table 2 highlighted in red for easy identification. We apologize for this minor oversight.

16. Figure 2: I do not believe that presenting the % difference adds any meaningful information when presenting the results. Consider revising this figure to present the change/difference in absolute points of the utility decrements. When doing this you may want to add information on statistical uncertainty (i.e. CI) in the figure to help interpret the level of (un)certainty in the results.
Reply：We sincerely appreciate your feedback on Figures 1 and 2, which has been instrumental in refining our data presentation.

To address your concerns:

For Figure 1 (QLU-C10D utility decrements for the cancer patient sample and the general population), we incorporated statistical uncertainty as you suggested. Standard errors for cancer patient utility decrement are visualized as error bars in Figure 1, with

	the description updated to reflect this: “The standard errors for cancer patients' utility decrement—visualized as error bars in Figure 1—occasionally exceed these observed differences, with such error bars reflecting statistical uncertainty.” This ensures readers can gauge the reliability of utility decrement comparisons. For Figure 2 (now presenting absolute differences in utility decrements), we revised the visualization to focus on absolute changes (replacing percentage differences) with the baseline set at 0, aligning with clinical interpretability. The updated text for Figure 2 reads: “Figure 2 compares utility decrements across severity levels from the cancer patient value set and the general population value set (with the baseline set at 0), highlighting absolute differences. Cancer patients showed significantly larger decrements across all dimensions, except for certain levels of role and social functioning and sleep disorders, and all levels of pain. The largest differences were observed at level 2, particularly in symptom dimensions like fatigue, lack of appetite, and bowel problems.” While Figure 2 does not include error bars, statistical uncertainty is addressed in Figure 1, and this relationship is clarified in the text to ensure readers contextualize results with appropriate rigor. This separation of visualizations balances the need to convey statistical uncertainty (Figure 1) and clinical relevance via absolute differences (Figure 2), avoiding overcomplication while addressing your key concerns. Thank you again for your thoughtful input—your guidance has strengthened the manuscript’s clarity and transparency. 17. Figure 3 and Table 3: As indicated above - please add information on the way these hypothetical health states were defined and also ordered. Information on descriptive system of the QLU-C10D would be valuable here. Reply：Thank you for your suggestion regarding Figure 3 and Table 3. We have supplemented relevant information as advised: In the penultimate paragraph of "2.5 Data analysis", we added: "Additionally, 1,000 hypothetical health states were generated by randomly combining the 4 levels across the 10 QLU-C10D dimensions. We employed simple random sampling to ensure each dimension’s levels were proportionally represented, with combinations spanning mild (e.g., most dimensions at level 1-2) to severe (e.g., majority at level 3-4), including the best (all level 1) and worst (all level 4) extremes. This sampling strategy provided comprehensive coverage for utility comparison between the two value sets." In the description of Fig. 3, we added: "ordered according to the values from the general population." These revisions clarify the definition and ordering of hypothetical health states, as well as details of the QLU-C10D descriptive system. We appreciate your careful guidance. 18. Page 9, line 8: You mention Bland-Altman plots and Figure G - I am not able to see any of this in the manuscript. Either delete this section or add the figure. Also keep consistent with numbering the figures numerically. Reply：As mentioned earlier, Figure G (along with other supplementary tables and figures labeled with letters, such as Table A, Figure B, etc.) was initially intended to be included in the supplementary materials. However, we have encountered consistent failures when attempting to upload these supplementary materials through the system, which is an oversight on our part for not addressing this issue more promptly. We sincerely apologize for the inconvenience and confusion this has caused. To ensure these materials are accessible, we have now included them directly in the main manuscript, positioned immediately after the references, and have revised the text to clearly indicate this location. Thank you for your understanding and patience as we work to resolve this issue, and we appreciate your feedback in helping us improve the completeness and accessibility of our submission. 19.Table 3: Please provide further clarification on this table. What are extreme

subgroups? What were ES calculated for? What is the incremental utility referring to? It is hard for me interpreting the results presented in this table.
Reply：Thank you for your feedback on Table 3. We sincerely appreciate your pointing out the necessity for clearer explanations to facilitate interpretation. As per your suggestion, we have added the following details in the last paragraph of "2.5 Data analysis" in the manuscript to clarify:

" Mean utility scores for each subgroup were compared using one-way ANOVA. "Extreme subgroups" here refer to the two ends of the stratified variables: for cancer stage, these were Stage I (earliest) and Stage IV (most advanced); for ECOG performance status, they were Grade 0 (fully active) and Grade 4 (completely disabled). Effect sizes (ES) quantified the clinical significance of differences between these extreme subgroups, calculated as the mean score difference divided by their pooled standard deviation [64]. Incremental utility was defined as the absolute difference in mean utility values between these extreme subgroups, reflecting the utility change associated with moving from the most severe to the least severe end of the stratified variable. "

These additions are aimed at clarifying the definitions of extreme subgroups, the calculation and purpose of ES, and the meaning of incremental utility, thereby making the results presented in Table 3 more interpretable. Once again, thank you for your guidance in enhancing the clarity of this section.

20.Page 9, line 58-59: you write "The cancer patient value set has a broader range and yields lower utility values, reflecting the higher priority cancer patients assign to QLU-C10Dattributes."My understanding: In its essence, this would mean that patients assign a higher preference to symptom reduction than survival time in the DCE tasks. Consider adding this to the discussion. It is somewhat half of the truth to say patients assign higher priority to QLU-C10D domains.
Reply：Thank you for your insightful comment on the interpretation of the cancer patient value set. Your suggestion to clarify the trade-off between symptom reduction and survival time in DCE tasks has helped strengthen our discussion.

We have revised the relevant sentence to:
"The cancer patient value set—with a broader range and lower utility values—reflects a key finding from DCE tasks: cancer patients assigned higher priority to improvements in QLU-C10D attributes than to survival time extension alone."

This revision explicitly highlights the preference for symptom-related attributes over survival time, addressing your observation and providing a more precise interpretation. We greatly appreciate your guidance in refining this critical point.

21.Consider adding this reference when referring to the QLU-C10D:
Simone Seyringer, Micha Johannes Pilz, Femke Jansen, Matthias Büttner, Madeleine T. King, Richard Norman, Georg Kemmler, Virginie Nerich, Bernhard Holzner, Andrew Bottomley, Eva M. Gamper,
Cancer-specific utility instrument for health economic evaluations: A synopsis of the EORTC QLU-C10D user manual and current validity evidence, European Journal of Cancer, Volume 217, 2025.
Reply：Thank you very much for your valuable suggestion regarding adding the reference related to the QLU-C10D. We have incorporated this reference into the manuscript as you advised, and it now appears in the citation [27-32] in the sentence: " The QLU-C10D, by contrast, has shown high sensitivity in detecting subtle health changes and distinguishing between health states among cancer patients [27-32]. "

We truly appreciate you bringing this article to our attention in a timely manner. It is extremely necessary for our research topic, and we believe it enhances the robustness and comprehensiveness of our discussion on the QLU-C10D.

Reviewer #2:
Dear Reviewer,

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

Thank you wholeheartedly for your meticulous review and insightful feedback on our manuscript. Your detailed comments—from the contextualization of PRO-based value sets in China to the nuances of experimental design, data quality control, and analytical rigor—reflect a profound expertise that has already helped us identify critical areas for improvement. We are particularly grateful for your recognition of the study's relevance to the discussion on patient-derived versus general population value sets, and we fully agree that enhancing methodological detail and expanding sensitivity analyses will significantly strengthen the work.

We also appreciate your thoughtful observations on the generalizability of our sample, the need to clarify survey termination criteria, and the importance of addressing scale heterogeneity and preference heterogeneity. These points have guided us to refine our approach, ensuring greater transparency and robustness in our findings.

Rest assured, we are committed to addressing each of your suggestions comprehensively in the revision. Below is our point-by-point response, with revised sections highlighted in red for easy reference.

Thank you again for your time, expertise, and dedication to elevating the quality of our research.

1.Introduction - Suggest speaking to whether value sets based on PROs other than the QLQ-C10D are available for cancer patients in China, to further contextualize the need for this study.
Reply：Thank you for your suggestion to enhance the introduction by contextualizing the availability of PRO-based value sets for cancer patients in China. We have incorporated relevant content to address this, clarifying the need for our study:

" In China, value sets based on general population preferences have been developed for several generic PBMs, including the EQ-5D [21-24], EQ-5D-Y [25], and SF-6D [26]. While these PBMs are widely applied in HTA, they are not cancer-specific and may not fully capture the symptom burden and HRQoL changes experienced by cancer patients. The QLU-C10D, by contrast, has shown high sensitivity in detecting subtle health changes and distinguishing between health states among cancer patients [27-32]."

Additionally, we have rewritten the fourth paragraph of the discussion to align with this context. These revisions aim to highlight the gap in cancer-specific PRO-based value sets in China, thereby strengthening the rationale for our research. We greatly appreciate your guidance in enriching the study's contextual foundation.

2.Methods, recruitment - Recruitment focused on patients who received care at two Grade tertiary public hospitals in Harbin, China. Suggest that the authors speak to how patient characteristics compared to the population characteristics of Chinese patients diagnosed with cancer and the potential implications of recruiting patients from only two hospitals in a single city in China for generalizability of results.
Reply：Thank you for your insightful comment regarding the representativeness of our sample and the generalizability of our findings. We appreciate your suggestion to compare our patient characteristics with the broader population of Chinese cancer patients.

As reflected in the revised Table 1 (Socio-demographic characteristics of the study samples), we have added a note stating: "No significant difference was found between the distribution in this study and the latest incidence rates reported by the National Cancer Center and the International Agency for Research on Cancer (IARC) (https://www.iarc.who.int/wp-content/uploads/2024/02/pr345_E.pdf)." Specifically, we observed no statistically significant difference in gender distribution ($\chi^2 = 0.343$, $p = 0.558$), which was a key demographic variable we prioritized during sampling.
We also aimed to achieve quota sampling for cancer types to enhance representativeness. Unfortunately, due to limited collaboration from certain clinical departments, we were unable to achieve statistical matching for cancer type distribution. For this reason, we did not report these comparisons, as we felt they would not accurately reflect population-level patterns.

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

Regarding the implications of recruiting from only two hospitals in Harbin for generalizability, we fully acknowledge this limitation, which is explicitly addressed in our discussion: " Firstly, patient recruitment was limited to two tertiary public hospitals in Harbin, China. While this approach ensured data quality and feasibility, it may limit the generalizability of our findings, as the recruited sample might not fully represent the broader cancer patient population across China in terms of socioeconomic, regional, and clinical diversity."

We hope this clarification addresses your concerns. Thank you again for highlighting this important aspect, which helps strengthen the contextual interpretation of our results.

3.Methods, design - Details on the experimental design are vague (adhered to 'optimal design theory'). Recommend clarifying exactly which approach was used to select the experimental design (e.g., D-optimal procedures) and whether the design considered the need to estimate interaction effects.
Reply：Thank you for your valuable suggestion to clarify the details of our experimental design. We appreciate your emphasis on specificity, as it enhances the transparency of our methodology.

As revised in the manuscript, we have now explicitly detailed the experimental design approach:

" This study employed the DCE-based QLU-C10Dvaluation protocol described by King et al. [33], widely used thereafter in valuation studies across numerous countries [33-45]. Patients completed 16 choice sets, selected from 960 combinations via a D-efficiency procedure (a key method under optimal design theory) that prioritizes maximizing the efficiency of main effect parameter estimates. Interaction effects were not included to balance survey feasibility and statistical precision, consistent with common practices [33-45]. "

We hope this clarification meets your expectations. Thank you again for helping strengthen the rigor of our methods section.

4. Methods, data quality control - The authors state that "with trained interviewers clarifying DCE tasks, assisting as needed, or terminating surveys when necessary." In which situations was terminating the survey deemed necessary? How often did this occur?
Reply：Thank you for your question regarding the circumstances and frequency of survey termination. We have supplemented the relevant content in the Methods section to address this:

" First, face-to-face interviews were conducted to optimize survey comprehension and interactivity: trained interviewers clarified DCE tasks, offered assistance as required, and halted surveys exclusively under specific circumstances—including inability to understand the survey interface, persistent failure to grasp task requirements (despite repeated explanations), or voluntary requests to withdraw (e.g., due to fatigue). These situations were deemed to directly compromise data validity. A total of 7 surveys (1.06% of the initial sample) were terminated, resulting in a final sample of 659 participants. The primary reasons for termination were fatigue (n=5) and persistent comprehension difficulties (n=2). Notably, none of these terminations were due to uncertainty or difficulty in making choices during the choice tasks. "

This addition details the specific situations that warranted survey termination and provides quantitative data on its occurrence, improving the transparency of our data quality control process. We appreciate your attention to this methodological detail, which helps strengthen the rigor of our study.

5. Methods, data quality control - The authors also state that they removed respondents who indicated that they found the survey difficult from the estimation. Please include the number of affected respondents. Past research suggests that DCE respondents who are uncertain in their responses may be engaging in more deliberative reasoning (DOI: 10.1016/j.socec.2014.02.009). Suggest performing sensitivity analysis on this exclusion criterion.

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

Reply：We appreciate your suggestion. It should be clarified that the 7 participants (accounting for 1.06% of the 659 finally included samples) whose surveys were terminated in this study were not excluded due to uncertainty or difficulty in making choices during the choice tasks. Instead, they were excluded because they were unable to understand the survey interface, persistently failed to grasp the task requirements (even after repeated explanations), or voluntarily requested to withdraw (e.g., due to fatigue)—situations that would directly compromise data validity.

Regarding sensitivity analysis, we acknowledge its importance in assessing result robustness. However, we did not perform a sensitivity analysis excluding these 7 participants for two reasons: First, their exclusion was based on invalidity (e.g., inability to complete tasks) rather than choice-related uncertainty, which is typically the focus of sensitivity analyses in DCE studies. Second, the small proportion (1.06%) makes it unlikely to meaningfully affect the overall results, as their exclusion would not alter the statistical patterns observed in the full sample.

That said, we recognize the value of such analyses in similar contexts. In future studies, we will refer to relevant literature (DOI: 10.1016/j.socec.2014.02.009) and include sensitivity analyses when excluding samples due to factors like choice uncertainty to further verify result robustness.

Thank you again for your thoughtful feedback.

6.Methods, data analysis - The authors state that QALYs were anchored to a 0 to scale, but in the results they report a negative value for the worst health state. Suggest clarifying whether results were scaled after initial model estimation or not, and if so, whether results sensitive to the scaling approach.
Reply：Thank you for your attention to our analytical approach for DCE data. We have refined the description of the utility estimation method to enhance clarity and transparency, as follows:

" Conditional logistic regression was employed to model utility and calculate utility weights by interacting QLU-C10D dimensions with survival time. Utility values were anchored at 0 ("dead") and 1 ("full health") to maintain ratio-scale properties[60, 61],yet remained unconstrained during initial estimation to capture states "worse than deaths"."
This approach aligns with methodological conventions in preference-based measurement, where unconstrained estimation of utilities (permitting values outside the 0-1 range) is often adopted to reflect the clinical reality that severe health states may be perceived as worse than death—a phenomenon particularly relevant in cancer populations. By retaining this flexibility, our analysis aims to capture the full spectrum of patient preferences, including extreme health states that generic value sets may overlook.

We believe this clarification better articulates the rationale behind our analytical choices, grounding them in established methods while emphasizing their relevance to our study population. Thank you again for prompting this refinement, which strengthens the methodological rigor of our work.

7.Method, data analysis - Suggest that the authors explore preference heterogeneity, through inclusion of interaction terms and/or mixed logit modelling. It is possible that different patient subgroups have different preferences for health state attributes and it would be informative to understand the level of variation within the sample.
Reply：Thank you profoundly for this invaluable suggestion to explore preference heterogeneity—your insight not only aligns perfectly with our own research goals but even resonates with the original intent behind how we designed our study sample. From the outset, we recognized that capturing the full complexity of health preferences would require examining both inter-group differences (between cancer patients and the general population, which is the focus of this manuscript) and intra-group variations within patient subgroups. In fact, exploring preference heterogeneity among cancer patients was explicitly the second core objective in designing our research sample. Your recommendation to delve into this heterogeneity feels like a serendipitous convergence of ideas, and we are genuinely excited by how it reinforces the depth of inquiry we aimed to achieve from the start.

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

To clarify, the current manuscript specifically focuses on comparing QLU-C10D preferences between Chinese cancer patients and the general population—a targeted analysis of inter-group differences. However, we fully agree that exploring subgroup variations within patient cohorts is equally critical, and this aligns seamlessly with your suggestion.

Notably, the companion study dedicated to this topic has already been completed, and it has been developed in parallel with the current manuscript. Using mixed logit modeling and interaction terms, it dissects preference heterogeneity among cancer patients, with the tentative title "Exploring Preference Heterogeneity in QLU-C10D Health State Valuations: Variations Across Cancer Types and Treatment Modalities." This study investigates how preferences for health attributes differ across subgroups (e.g., by cancer type, age, and treatment modality).

Preliminary findings include, for example, that younger breast cancer patients prioritize "emotional function" more than older patients, while elderly lung cancer patients value "physical function" higher. Patients on immunotherapy also place greater weight on "skin symptoms" compared to those receiving chemotherapy—likely due to distinct adverse effect profiles. These insights complement the current manuscript, and we plan to reference them in future discussions to provide a more comprehensive picture.

Thank you again for this astute suggestion, which not only enhances our work but also validates the foundational vision we had for this research. We truly look forward to the possibility of you encountering this companion study on heterogeneity in the future—it would mean a great deal to us to have your perspective on it.

8. Method, data analysis - The authors state that "Non-monotonic coefficients were adjusted by constraining adjacent levels, as per standard practice." The choice to constrain attribute levels should be supported by the results of likelihood-ratio testing indicating that the part-worth utility values for the levels are not significantly different from one another. Suggest reporting on these hypothesis test results.
Reply：Thank you profoundly for this insightful suggestion, which has given us a new perspective we hadn’t previously considered. Your point about validating the constraint of adjacent levels through likelihood-ratio testing is crucial for enhancing the rigor of our methodology, and we truly appreciate you highlighting this.

Following your guidance, we conducted supplementary analyses in Stata, specifically performing a likelihood-ratio test on the non-monotonicity observed between role functioning levels 3 and 4—with level 3 as the reference. The results showed no statistically significant difference between these two levels ($p = 0.21$), confirming that participants’ preferences for them are not distinct. This validates the rationale for imposing monotonicity constraints to resolve the inconsistency.

We have accordingly added the following to the manuscript: "A likelihood-ratio test confirmed no significant difference between these levels, so monotonicity was imposed and the model re-estimated."

What’s more, this approach to validating constraints through statistical testing is likely to inform many of our future studies, as it sets a stronger standard for methodological rigor. Thank you again for this invaluable input, which has significantly strengthened our work.

9. Method, data analysis - Suggest including how the utility value for any health state was calculated based on the estimated utility decrements for transparency.
Reply：Thank you for suggesting that we clarify how utility values for health states are calculated based on estimated utility decrements, to enhance transparency. Your suggestion has made us realize that this detail is indeed crucial for ensuring the methodology is understandable and reproducible.

We have added a specific example in the last part of "3.3 Utility estimates" to illustrate this process:

" For example, for QLQ-C30 responses matching the QLU-C10D health state

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

	"3233412432," the utility score is calculated as: 1 - 0.275 - 0.052 - 0.035 - 0.057 - 0.125 - 0 - 0.046 - 0.083 - 0.075 - 0.030 = 0.222. The worst possible health state (PITS state 4444444444) had a utility value of −0.092, considerably lower than the Chinese general population PITS value (0.083)[45]." This example, provided after establishing the value set, aims to clearly demonstrate the calculation logic. We appreciate your guidance in improving the clarity of our methodology. 10. Discussion - This study compares results from non-linear analyses across two populations. Suggest speaking to scale heterogeneity in this setting (DOI: 10.1007/s40271-017-0282-4) and comparisons when models upon which utility decrements were based included different underlying parameters (given differences in non-monotonicity observed across populations varied and required different constraints). Reply：Thank you for highlighting the need to address scale heterogeneity and model parameter differences in our cross-population comparison. Your reference to Rose et al. (2017) has been incredibly illuminating—frankly, your consecutive suggestions have opened up fascinating new angles we hadn't fully explored, making the study feel richer and more layered with each insight. It's exciting to see how much more there is to unpack! As suggested, we've added a concise note in the discussion to acknowledge these factors: " It is important to note that cross-population differences could also reflect scale heterogeneity (i.e., variations in choice consistency) [70], as well as differences in model constraints applied to address non-monotonicities between groups. These factors may influence utility estimates beyond true preference differences, warranting cautious interpretation." This revision frames potential influences on our findings without overcomplicating the discussion, as you recommended. Thank you again for these eye-opening inputs—they've not only strengthened our interpretation but also sparked new curiosity about deeper layers we can explore. 11. Discussion - This sentence should include a reference for global incidence numbers "This study developed a QLU-C10D value set based on the health preferences of cancer patients in China, the country with the highest cancer incidence globally, and compared the...

Dear Editor,

We sincerely appreciate the editor’s kind attention and voluntary work in handling our manuscript. Thank you for your time and support.

Reviewers' comments:

Reviewer #1:

Dear Reviewer：

Thank you from the bottom of our hearts for your extraordinarily thorough and insightful review of our manuscript. Your attention to even the most detailed aspects—including the significance of statistical coefficients and the clarity of figures—reflects a level of rigor that we deeply admire and appreciate. We are particularly grateful for your recognition of our efforts in recruiting a large number of cancer patients within a short timeframe; this was indeed a challenging feat, as we had only a limited window of approval from the participating hospitals. To meet this timeline, our team mobilized significant human resources—including trained interviewers, clinical collaborators, and data managers—working closely to streamline recruitment, ensure participant comfort, and maintain data quality. Your acknowledgment of this effort means a great deal to us.

We wholeheartedly agree with your observations: there were indeed points in our manuscript that lacked sufficient clarification, and revisions to the figures are absolutely necessary to enhance readability and precision. Your feedback has pinpointed exactly where we need to strengthen the work, and we are committed to addressing every single one of your suggestions with the utmost care.

Rest assured, we will revise the manuscript diligently based on your expert guidance—clarifying ambiguous details, refining the figures to better convey key findings, and ensuring all analyses are presented with the rigor they deserve. We believe this process will significantly elevate the quality of our work, and we owe this to your thoughtful input.

Thank you again for your time, expertise, and dedication to helping us improve. Below is our point-by-point response, with revised sections highlighted in red for easy reference.

1. The authors need to be very careful with using the terminology on the concepts presented correctly and coherently. I suggest using the term "value sets" for anything that refers to the utility decrements of the patients or general populations. Be careful to not confuse the terms QALYs, utility values, value sets, utility decrements, and preferences. The field needs to be very careful in using these terms correctly as these concepts are all similar but have a critical distinction.

Reply： Thank you profoundly for this critical and insightful feedback. We fully agree that precise and consistent use of terminology is foundational to clarity in health economics and preference research, and we deeply appreciate you highlighting this important point—your guidance will significantly enhance the rigor of our manuscript.

We recognize that our initial use of terms such as “value sets,” “QALYs,” “utility values,” “utility decrements,” and “preferences” may not have consistently reflected their nuanced distinctions, and we take this concern very seriously. To address this, we have conducted a systematic review of the entire manuscript (including text, figures, tables, and references) to ensure each term is used with strict adherence to its established definition in the field:

“Value sets” will specifically refer to the preference-based scoring algorithms derived from the respective populations (cancer patients and the general population) used to quantify health states.
“Utility values” will denote the single-index scores assigned to health states (e.g., from 0 to 1) using these value sets.
“Utility decrements” will describe the differences between these values (e.g., from a reference state to an impaired state), capturing the magnitude of health-related quality of life loss.
“QALYs” will be reserved for the metric integrating utility values over time, reflecting the overall health benefit of an intervention.
“Preferences” will refer to the underlying judgments (e.g., trade-offs) guiding these valuations, as elicited through methods like discrete choice experiments.

Notably, we have revised the titles of key figures to align with these definitions, moving away from previous references to “utility decrements.” Specifically:

Fig. 1 is now titled “QLU-C10D value sets for the cancer patient sample and the general population,” directly focusing on the scoring algorithms themselves.
Fig. 2 has been updated to “Difference in value sets between cancer patients and the general population,” explicitly emphasizing the comparison of these algorithmic frameworks.

These adjustments extend to the entire text, where descriptions of analyses and results are strictly aligned with the defined terms. All revisions—whether in figures, text, or tables—are marked in red for easy reference.

Thank you again for prompting us to refine these critical distinctions—your expertise has been invaluable in elevating the quality of our work.

2. One of the main conclusions the authors wish to draw is that patients' value sets lead to lower utility values compared to the value sets of the general population (this is often stated). While I understand that this claim is supported by applying the value sets on

hypothetical health states of the QLU-C10D descriptive system, it lacks to inform the reader how these hypothetical health states were defined. As the authors present this as the main finding, detailed information on the definition of the hypothetical health states is required to inform the readership.

Reply： Thank you sincerely for highlighting this critical point. We fully agree that detailed information on the definition of hypothetical health states is essential to contextualizing our main conclusion, and we appreciate your guidance in strengthening the transparency of our analysis.

As suggested, we have supplemented the third paragraph of Section 2.5 (Data Analysis) with specific details on the hypothetical health states (revisions marked in red):

" Additionally, 1,000 hypothetical health states were generated by randomly combining the 4 levels across the 10 dimensions of the QLU-C10D descriptive system. This sampling approach ensured proportional representation of each dimension's levels, with states spanning the full severity spectrum—from mild (e.g., most dimensions at levels 1–2) to severe (e.g., majority at levels 3–4)—and explicitly including the best-possible (all dimensions at level 1) and worst-possible (all dimensions at level 4) health states. This comprehensive coverage of potential health states was designed to ensure robust utility comparisons between the two value sets."

We believe these details adequately clarify how the hypothetical health states were defined and justify the generalizability of our findings. Thank you again for prompting this important improvement—your input has significantly enhanced the rigor of our methodology.

3.The general population value sets for mainland China were weighted for education and "geography" (I believe to achieve representation of persons in different provinces). This study does not weight the results. I understand that it is not possible to do so for "geography" but would be possible for the variable "education". I would recommend doing so (to reflect a similar distribution as in the gen. pop. publication), if the authors wish to abstain from that I would suggest adding this in the limitation section.

Reply： Thank you sincerely for this thoughtful suggestion regarding education weighting—your insight has helped us reflect more deeply on the alignment between our methodology and the general population value set, and we truly appreciate you pushing us to clarify this important methodological choice.

We completely agree that education weighting is a valuable practice for general population studies, and we took your recommendation very seriously: in fact, we explicitly tested this approach during our pre-analysis phase (as you'll see in our revised limitations section) to ensure we weren't overlooking a critical step.

As we worked through the analysis, we faced a few key considerations that led us to prioritize unweighted results, and we'd like to share them with you in the spirit of transparency:

First, the demographic profile of cancer patients in China—particularly age and education distributions—differs inherently from the general population (e.g., with higher representation in older age groups, who often have distinct educational backgrounds). Unfortunately, there's a lack of robust, disease-specific demographic benchmarks for China's cancer patient population, making it challenging to derive weighting criteria that truly reflect the realities of this group.

Second, our sampling strategy focused on high-incidence cancer types to ensure our findings would resonate with the most common clinical scenarios in China's oncology practice. In preliminary tests, we found that forcing alignment with general population education distributions risked diluting the unique preferences of these patients—especially differences in how they value health states across the cancer types we prioritized. Since our core goal was to capture patient-specific value sets, we worried this might undermine the ecological validity of our work.

We wholeheartedly agree that this is a limitation, and we've added a detailed note to the limitations section (marked in red) to acknowledge it: "Thirdly, we did not apply education weighting (as used in the general population value set) due to inherent differences in education distributions between cancer patients and the general population; this omission may weaken the stability of extrapolating our findings for cross-group comparisons."

Importantly, while we opted not to apply education weighting in this manuscript, we fully recognize the significance of educational background in shaping health preferences—and this is precisely why we've dedicated a companion study to exploring such nuances in depth. Developed in parallel with the current work, this companion manuscript focuses on unpacking preference heterogeneity among cancer patients, with a specific emphasis on how educational attainment intersects with other factors (e.g., cancer type, age, and treatment modality) to influence health state valuations.

Tentatively titled "Exploring Preference Heterogeneity in QLU-C10D Health State Valuations: Variations Across Cancer Types and Treatment Modalities," this study uses mixed logit modeling and interaction terms to dissect subgroup differences. For instance, preliminary analyses reveal that within the same cancer type, patients with different educational backgrounds may prioritize certain health attributes differently—such as greater emphasis on "social function" among those with higher education, compared to stronger focus on "physical comfort" among those with lower education. These insights build on the findings of the current manuscript, offering a more granular understanding of how education, alongside other variables, shapes patient preferences.

This companion study aligns closely with your suggestion to explore the role of education, and we believe it provides the depth of analysis this topic deserves. We plan to reference its key findings in the discussion section of the current manuscript to contextualize our unweighted results and highlight avenues for future research.

Thank you again for prompting this careful consideration—your attention to this detail has strengthened the rigor of our methodology, and we're grateful for the opportunity to clarify our approach. We hope this explanation addresses your concern, and we'd be delighted to share more details about the companion study if you're interested. Please feel free to reach out for further discussion!

4.The manuscript would benefit from a proofreading service. This would help to align the consistency of the statements made and help with the use of the terminology as well.
Reply： Thank you for your suggestion regarding proofreading. We appreciate your attention to the consistency of statements and terminology.

Notably, several authors on our team are native English speakers, and they have further revised and polished the manuscript on the basis of previous modifications to ensure linguistic accuracy, terminological consistency, and overall coherence.

We believe these efforts have enhanced the quality of the manuscript, and we thank you again for your valuable feedback.

5.The interpretation of the results should be further done taking measures of statistical uncertainty into account. The standard errors presented in the tables sometimes exceed the difference of the utility decrements of the gen. population and the patients. This is not taken into account when presenting the results.
Reply： We sincerely appreciate your insightful feedback on incorporating statistical uncertainty into result interpretation.

To address your concern, we have revised our analysis and visualization to explicitly account for statistical uncertainty:

We calculated the standard errors for cancer patient utility decrement and integrated them into Figure 1 as error bars. These error bars directly reflect statistical variability, and we have updated the textual description of the figure to clarify this:" The standard errors for cancer patients' utility decrement—visualized as error bars in Figure 1—occasionally exceed these observed differences, with such error bars reflecting statistical uncertainty."

By visualizing standard errors, we ensure that readers can now interpret utility decrement differences (e.g., cases where standard errors surpass these differences) with full awareness of statistical uncertainty. This addition helps contextualize the reliability

of our findings and aligns with your recommendation to consider precision in result interpretation.

Abstract:

6. Page 1, line 29-30: "For China, utility scores based on the general population value set underestimate quality adjusted life years (QALY) compared to those based on the cancer patient value set." This conclusion is not supported by the analysis. 1.) you did not calculate QALYs, 2.) utility values are higher in the gen. pop. Potentially leading to an overestimation of QALYs.

Reply： Thank you for your astute observation regarding our inappropriate conclusion about QALYs. We fully agree with your points: we did not calculate QALYs in this study, and the higher utility values in the general population value set could potentially lead to overestimation of QALYs, making our original claim unsupported.

To address this, we have revised the problematic statement to strictly align with our analysis, which focuses on utility value sets rather than QALY calculations. The updated conclusion now reads:

"Differences in value sets between cancer patients and the general population underscore the importance of more comprehensively integrating patient perspectives into existing evaluation frameworks."

This revision removes all references to QALYs, anchors the conclusion in our core findings (comparisons of utility value set properties), and emphasizes the practical relevance to cost-effectiveness analyses—consistent with the scope and results of our study.

Thank you again for guiding us to refine this conclusion, enhancing the accuracy and rigor of our work.

Introduction:

7. Page 2, line 8: Consider revising "HRQoL utility values" to "Country specific value sets, reflecting societal preferences"

Reply： Thank you for your suggestion to revise "HRQoL utility values". We have followed your advice and revised the phrase to "Country specific value sets, reflecting societal preferences" as recommended. This change aligns with your insight and improves the precision of the statement. Thank you again for your valuable input.

8. Page 2, line 17: consider revising "patient preferences" to "value sets represent the preferences informed by the lived experience".

Reply： We have adopted it and revised the content to: "Therefore, integrating value sets representing preferences informed by the lived experience alongside societal

perspectives offers a more comprehensive framework for clinical and policy decision-making." Your guidance helps improve the precision of our expression.

9. Page 2, line 19: I do not believe that Germany relies on QALY assessment at all. Read IQWIG guidelines for a proper understanding of the German HTA strategies.
Reply： Following your advice, we have revised the relevant text to: "In Germany, the HTA framework led by the Institute for Quality and Efficiency in Health Care (IQWiG) focuses on qualitative assessments of 'added clinical benefit' from the patient perspective, such as survival gains and symptom improvement. It directly incorporates patient-reported outcomes (PROs) to evaluate therapies against comparators, ensuring patient voices are central to treatment benefit evaluations."

This revision removes any implication of QALY reliance and aligns with the actual focus of Germany's HTA strategy, as highlighted in your comment. Thank you again for helping us improve the accuracy of our work.

10. Further I want to provide an excerpt from the Swedish value set publication: "Sweden relies on cost-effectiveness analysis in health economic appraisal, but this analysis is subordinate to the other two guiding principles of human value and need and solidarity when making resource-allocation decisions [1, 4]. In general, the use of CUA and QALYs is recommended by the Dental and Pharmaceutical Benefits Agency [41]."
Reply： Thank you for providing the excerpt from the Swedish value set publication, which offers valuable context for Sweden's health economic appraisal approach. We have incorporated this insight while retaining our focus on patient-centered elements in the revised text:

" Similarly, Sweden's Dental and Pharmaceutical Benefits Agency (TLV) emphasizes ethical principles like "human value and need" and "solidarity" in its health economic evaluations, which prioritize patient experiences. These evaluations integrate condition-specific PROs to align resource allocation with patients' actual needs [8]."

This revision stays true to the core of our focus on patient preferences and the role of ethical principles in TLV's framework, as highlighted in your reference. We appreciate your guidance in ensuring the accuracy of this section.

Additional Minor Points:
11. Page 2, line 37-38: English: "of universally applying…."
Reply： We have revised the text to "of universally applying…" as suggested, which aligns better with standard English usage. Your careful review is greatly appreciated, as it contributes significantly to enhancing the precision and professionalism of our manuscript.

12.Page 2, line 44-45: Correct full name: Quality of Life Questionnaire - Core 30

Reply： Thank you for correcting the name "Quality of Life Questionnaire - Core 30" on Page 2, lines 44-45. We apologize for this oversight in our field and have revised it as advised. Your careful review is much appreciated.

13.Page 3, line 1: Here and in the discussion you write: "with the highest global cancer incidence". If this is true, please provide a reference to the statement (e.g. GLOBOCAN, IARC data).

Reply： As advised, we have supplemented the relevant parts with a specific reference to validate the statement: Han B, Zheng R, Zeng H, Wang S, Sun K, Chen R, Li L, Wei W, He J: Cancer incidence and mortality in China, 2022. J Natl Cancer Cent 2024, 4(1):47-53. Specifically, this reference has been added and highlighted in red both in the corresponding part of the introduction and in the first paragraph of the discussion where the same statement appears.

Valuation Survey and DCE Design:

14. Page 4, line 7: you write "five of ten HRQoL dimensions" - typically four HRQoL dimensions + survival time. Also the dimension "physical functioning" is presented as two items in the conventional DCE tasks (displaying 10 dimensions, but 11 items). Please correct for accurate writing.

Reply： Thank you for your careful observation regarding the description of HRQoL dimensions and attributes in the DCE tasks. We appreciate you pointing out the need to clarify that the conventional design typically includes four HRQoL dimensions plus survival time, as well as the distinction between the 10 domains (with "physical functioning" split into 2 attributes, totaling 11 items).

We have revised the text to accurately reflect these details:

" To control biases, Health State A/B assignment and the order of 11 attributes across 10 QLU-C10D domains (the physical functioning dimension was split into 2 attributes, one for ‘long walk’ and one for ‘short walk’ (see Table A) were randomized per respondent; a balanced incomplete block design (BIBD) defined four differing dimensions per set (others held equal) to minimize burden. Each set presented two hypothetical QLU-C10D health states, differing in four of ten HRQoL dimensions and survival times (1, 2, 5, or 10 years), with differences highlighted for clarity. Examples are shown in Figures A and B (Chinese version) in the supplementary material. Details on QLU-C10D methodology, including formats, attribute ordering, and reliability, are provided in related studies [10, 57-59]."

This revision aims to align with the conventional structure of DCE tasks and accurately clarify the relationship between domains, attributes, and dimensions. Thank you again for helping us refine the precision of this description.

Results:

15. Table 2: You fail to indicate which coefficients are statistically significant. This is indicated in the footnote of the table, not a single coefficient is marked with an .

Reply： We have added the relevant markers as advised, with the revised sections in Table 2 highlighted in red for easy identification. We apologize for this minor oversight.

16. Figure 2: I do not believe that presenting the % difference adds any meaningful information when presenting the results. Consider revising this figure to present the change/difference in absolute points of the utility decrements. When doing this you may want to add information on statistical uncertainty (i.e. CI) in the figure to help interpret the level of (un)certainty in the results.

Reply： We sincerely appreciate your feedback on Figures 1 and 2, which has been instrumental in refining our data presentation.

To address your concerns:

For Figure 1 (QLU-C10D utility decrements for the cancer patient sample and the general population), we incorporated statistical uncertainty as you suggested. Standard errors for cancer patient utility decrement are visualized as error bars in Figure 1, with the description updated to reflect this: “The standard errors for cancer patients' utility decrement—visualized as error bars in Figure 1—occasionally exceed these observed differences, with such error bars reflecting statistical uncertainty.” This ensures readers can gauge the reliability of utility decrement comparisons.

For Figure 2 (now presenting absolute differences in utility decrements), we revised the visualization to focus on absolute changes (replacing percentage differences) with the baseline set at 0, aligning with clinical interpretability. The updated text for Figure 2 reads: “Figure 2 compares utility decrements across severity levels from the cancer patient value set and the general population value set (with the baseline set at 0), highlighting absolute differences. Cancer patients showed significantly larger decrements across all dimensions, except for certain levels of role and social functioning and sleep disorders, and all levels of pain. The largest differences were observed at level 2, particularly in symptom dimensions like fatigue, lack of appetite, and bowel problems.” While Figure 2 does not include error bars, statistical uncertainty is addressed in Figure 1, and this relationship is clarified in the text to ensure readers contextualize results with appropriate rigor.

This separation of visualizations balances the need to convey statistical uncertainty (Figure 1) and clinical relevance via absolute differences (Figure 2), avoiding overcomplication while addressing your key concerns.

Thank you again for your thoughtful input—your guidance has strengthened the manuscript’s clarity and transparency.

17. Figure 3 and Table 3: As indicated above - please add information on the way these hypothetical health states were defined and also ordered. Information on descriptive system of the QLU-C10D would be valuable here.

Reply： Thank you for your suggestion regarding Figure 3 and Table 3. We have supplemented relevant information as advised:

In the penultimate paragraph of "2.5 Data analysis", we added: "Additionally, 1,000 hypothetical health states were generated by randomly combining the 4 levels across the 10 QLU-C10D dimensions. We employed simple random sampling to ensure each dimension's levels were proportionally represented, with combinations spanning mild (e.g., most dimensions at level 1-2) to severe (e.g., majority at level 3-4), including the best (all level 1) and worst (all level 4) extremes. This sampling strategy provided comprehensive coverage for utility comparison between the two value sets."

In the description of Fig. 3, we added: "ordered according to the values from the general population."

These revisions clarify the definition and ordering of hypothetical health states, as well as details of the QLU-C10D descriptive system. We appreciate your careful guidance.

18. Page 9, line 8: You mention Bland-Altman plots and Figure G - I am not able to see any of this in the manuscript. Either delete this section or add the figure. Also keep consistent with numbering the figures numerically.

Reply： As mentioned earlier, Figure G (along with other supplementary tables and figures labeled with letters, such as Table A, Figure B, etc.) was initially intended to be included in the supplementary materials. However, we have encountered consistent failures when attempting to upload these supplementary materials through the system, which is an oversight on our part for not addressing this issue more promptly.

We sincerely apologize for the inconvenience and confusion this has caused. To ensure these materials are accessible, we have now included them directly in the main manuscript, positioned immediately after the references, and have revised the text to clearly indicate this location.

Thank you for your understanding and patience as we work to resolve this issue, and we appreciate your feedback in helping us improve the completeness and accessibility of our submission.

19.Table 3: Please provide further clarification on this table. What are extreme subgroups? What were ES calculated for? What is the incremental utility referring to? It is hard for me interpreting the results presented in this table.

Reply： Thank you for your feedback on Table 3. We sincerely appreciate your pointing out the necessity for clearer explanations to facilitate interpretation. As per

your suggestion, we have added the following details in the last paragraph of "2.5 Data analysis" in the manuscript to clarify:

" Mean utility scores for each subgroup were compared using one-way ANOVA. "Extreme subgroups" here refer to the two ends of the stratified variables: for cancer stage, these were Stage I (earliest) and Stage IV (most advanced); for ECOG performance status, they were Grade 0 (fully active) and Grade 4 (completely disabled). Effect sizes (ES) quantified the clinical significance of differences between these extreme subgroups, calculated as the mean score difference divided by their pooled standard deviation [64]. Incremental utility was defined as the absolute difference in mean utility values between these extreme subgroups, reflecting the utility change associated with moving from the most severe to the least severe end of the stratified variable. "

These additions are aimed at clarifying the definitions of extreme subgroups, the calculation and purpose of ES, and the meaning of incremental utility, thereby making the results presented in Table 3 more interpretable. Once again, thank you for your guidance in enhancing the clarity of this section.

20.Page 9, line 58-59: you write "The cancer patient value set has a broader range and yields lower utility values, reflecting the higher priority cancer patients assign to QLU-C10Dattributes."My understanding: In its essence, this would mean that patients assign a higher preference to symptom reduction than survival time in the DCE tasks. Consider adding this to the discussion. It is somewhat half of the truth to say patients assign higher priority to QLU-C10D domains.

Reply： Thank you for your insightful comment on the interpretation of the cancer patient value set. Your suggestion to clarify the trade-off between symptom reduction and survival time in DCE tasks has helped strengthen our discussion.

We have revised the relevant sentence to:
"The cancer patient value set—with a broader range and lower utility values—reflects a key finding from DCE tasks: cancer patients assigned higher priority to improvements in QLU-C10D attributes than to survival time extension alone."

This revision explicitly highlights the preference for symptom-related attributes over survival time, addressing your observation and providing a more precise interpretation. We greatly appreciate your guidance in refining this critical point.

21.Consider adding this reference when referring to the QLU-C10D: Simone Seyringer, Micha Johannes Pilz, Femke Jansen, Matthias Büttner, Madeleine T. King, Richard Norman, Georg Kemmler, Virginie Nerich, Bernhard Holzner, Andrew Bottomley, Eva M. Gamper, Cancer-specific utility instrument for health economic evaluations: A synopsis of the

EORTC QLU-C10D user manual and current validity evidence, European Journal of Cancer, Volume 217, 2025.

Reply： Thank you very much for your valuable suggestion regarding adding the reference related to the QLU-C10D. We have incorporated this reference into the manuscript as you advised, and it now appears in the citation [27-32] in the sentence: " The QLU-C10D, by contrast, has shown high sensitivity in detecting subtle health changes and distinguishing between health states among cancer patients [27-32]. "

We truly appreciate you bringing this article to our attention in a timely manner. It is extremely necessary for our research topic, and we believe it enhances the robustness and comprehensiveness of our discussion on the QLU-C10D.

Reviewer #2:
Dear Reviewer,

Thank you wholeheartedly for your meticulous review and insightful feedback on our manuscript. Your detailed comments—from the contextualization of PRO-based value sets in China to the nuances of experimental design, data quality control, and analytical rigor—reflect a profound expertise that has already helped us identify critical areas for improvement. We are particularly grateful for your recognition of the study's relevance to the discussion on patient-derived versus general population value sets, and we fully agree that enhancing methodological detail and expanding sensitivity analyses will significantly strengthen the work.

We also appreciate your thoughtful observations on the generalizability of our sample, the need to clarify survey termination criteria, and the importance of addressing scale heterogeneity and preference heterogeneity. These points have guided us to refine our approach, ensuring greater transparency and robustness in our findings.

Rest assured, we are committed to addressing each of your suggestions comprehensively in the revision. Below is our point-by-point response, with revised sections highlighted in red for easy reference.

Thank you again for your time, expertise, and dedication to elevating the quality of our research.

1.Introduction - Suggest speaking to whether value sets based on PROs other than the QLQ-C10D are available for cancer patients in China, to further contextualize the need for this study.

Reply： Thank you for your suggestion to enhance the introduction by contextualizing the availability of PRO-based value sets for cancer patients in China. We have incorporated relevant content to address this, clarifying the need for our study:

" In China, value sets based on general population preferences have been developed for several generic PBMs, including the EQ-5D [21-24], EQ-5D-Y [25], and SF-6D [26]. While these PBMs are widely applied in HTA, they are not cancer-specific and may not fully capture the symptom burden and HRQoL changes experienced by cancer patients. The QLU-C10D, by contrast, has shown high sensitivity in detecting subtle health changes and distinguishing between health states among cancer patients [27-32]."

Additionally, we have rewritten the fourth paragraph of the discussion to align with this context. These revisions aim to highlight the gap in cancer-specific PRO-based value sets in China, thereby strengthening the rationale for our research. We greatly appreciate your guidance in enriching the study's contextual foundation.

2.Methods, recruitment - Recruitment focused on patients who received care at two Grade tertiary public hospitals in Harbin, China. Suggest that the authors speak to how patient characteristics compared to the population characteristics of Chinese patients diagnosed with cancer and the potential implications of recruiting patients from only two hospitals in a single city in China for generalizability of results.

Reply： Thank you for your insightful comment regarding the representativeness of our sample and the generalizability of our findings. We appreciate your suggestion to compare our patient characteristics with the broader population of Chinese cancer patients.

As reflected in the revised Table 1 (Socio-demographic characteristics of the study samples), we have added a note stating: "No significant difference was found between the distribution in this study and the latest incidence rates reported by the National Cancer Center and the International Agency for Research on Cancer (IARC) (https://www.iarc.who.int/wp-content/uploads/2024/02/pr345_E.pdf)." Specifically, we observed no statistically significant difference in gender distribution ($\chi^2 = 0.343$, $p = 0.558$), which was a key demographic variable we prioritized during sampling.

We also aimed to achieve quota sampling for cancer types to enhance representativeness. Unfortunately, due to limited collaboration from certain clinical departments, we were unable to achieve statistical matching for cancer type distribution. For this reason, we did not report these comparisons, as we felt they would not accurately reflect population-level patterns.

Regarding the implications of recruiting from only two hospitals in Harbin for generalizability, we fully acknowledge this limitation, which is explicitly addressed in our discussion: " Firstly, patient recruitment was limited to two tertiary public hospitals in Harbin, China. While this approach ensured data quality and feasibility, it may limit the generalizability of our findings, as the recruited sample might not fully represent the broader cancer patient population across China in terms of socioeconomic, regional, and clinical diversity."

We hope this clarification addresses your concerns. Thank you again for highlighting this important aspect, which helps strengthen the contextual interpretation of our results.

3.Methods, design - Details on the experimental design are vague (adhered to 'optimal design theory'). Recommend clarifying exactly which approach was used to select the experimental design (e.g., D-optimal procedures) and whether the design considered the need to estimate interaction effects.

Reply： Thank you for your valuable suggestion to clarify the details of our experimental design. We appreciate your emphasis on specificity, as it enhances the transparency of our methodology.

As revised in the manuscript, we have now explicitly detailed the experimental design approach:

" This study employed the DCE-based QLU-C10Dvaluation protocol described by King et al. [33], widely used thereafter in valuation studies across numerous countries [33-45]. Patients completed 16 choice sets, selected from 960 combinations via a D-efficiency procedure (a key method under optimal design theory) that prioritizes maximizing the efficiency of main effect parameter estimates. Interaction effects were not included to balance survey feasibility and statistical precision, consistent with common practices [33-45]. "

We hope this clarification meets your expectations. Thank you again for helping strengthen the rigor of our methods section.

4. Methods, data quality control - The authors state that "with trained interviewers clarifying DCE tasks, assisting as needed, or terminating surveys when necessary." In which situations was terminating the survey deemed necessary? How often did this occur?

Reply： Thank you for your question regarding the circumstances and frequency of survey termination. We have supplemented the relevant content in the Methods section to address this:

" First, face-to-face interviews were conducted to optimize survey comprehension and interactivity: trained interviewers clarified DCE tasks, offered assistance as required, and halted surveys exclusively under specific circumstances—including inability to understand the survey interface, persistent failure to grasp task requirements (despite repeated explanations), or voluntary requests to withdraw (e.g., due to fatigue). These situations were deemed to directly compromise data validity. A total of 7 surveys (1.06% of the initial sample) were terminated, resulting in a final sample of 659 participants. The primary reasons for termination were fatigue (n=5) and persistent comprehension difficulties (n=2). Notably, none of these terminations were due to uncertainty or difficulty in making choices during the choice tasks. "

This addition details the specific situations that warranted survey termination and provides quantitative data on its occurrence, improving the transparency of our data quality control process. We appreciate your attention to this methodological detail, which helps strengthen the rigor of our study.

5. Methods, data quality control - The authors also state that they removed respondents who indicated that they found the survey difficult from the estimation. Please include the number of affected respondents. Past research suggests that DCE respondents who are uncertain in their responses may be engaging in more deliberative reasoning (DOI: 10.1016/j.socec.2014.02.009). Suggest performing sensitivity analysis on this exclusion criterion.

Reply： We appreciate your suggestion. It should be clarified that the 7 participants (accounting for 1.06% of the 659 finally included samples) whose surveys were

terminated in this study were not excluded due to uncertainty or difficulty in making choices during the choice tasks. Instead, they were excluded because they were unable to understand the survey interface, persistently failed to grasp the task requirements (even after repeated explanations), or voluntarily requested to withdraw (e.g., due to fatigue)—situations that would directly compromise data validity.

Regarding sensitivity analysis, we acknowledge its importance in assessing result robustness. However, we did not perform a sensitivity analysis excluding these 7 participants for two reasons: First, their exclusion was based on invalidity (e.g., inability to complete tasks) rather than choice-related uncertainty, which is typically the focus of sensitivity analyses in DCE studies. Second, the small proportion (1.06%) makes it unlikely to meaningfully affect the overall results, as their exclusion would not alter the statistical patterns observed in the full sample.

That said, we recognize the value of such analyses in similar contexts. In future studies, we will refer to relevant literature (DOI: 10.1016/j.socec.2014.02.009) and include sensitivity analyses when excluding samples due to factors like choice uncertainty to further verify result robustness.

Thank you again for your thoughtful feedback.

6.Methods, data analysis - The authors state that QALYs were anchored to a 0 to scale, but in the results they report a negative value for the worst health state. Suggest clarifying whether results were scaled after initial model estimation or not, and if so, whether results sensitive to the scaling approach.

Reply： Thank you for your attention to our analytical approach for DCE data. We have refined the description of the utility estimation method to enhance clarity and transparency, as follows:

" Conditional logistic regression was employed to model utility and calculate utility weights by interacting QLU-C10D dimensions with survival time. Utility values were anchored at 0 ("dead") and 1 ("full health") to maintain ratio-scale properties[60, 61],yet remained unconstrained during initial estimation to capture states "worse than deaths"." This approach aligns with methodological conventions in preference-based measurement, where unconstrained estimation of utilities (permitting values outside the 0-1 range) is often adopted to reflect the clinical reality that severe health states may be perceived as worse than death—a phenomenon particularly relevant in cancer populations. By retaining this flexibility, our analysis aims to capture the full spectrum of patient preferences, including extreme health states that generic value sets may overlook.

We believe this clarification better articulates the rationale behind our analytical choices, grounding them in established methods while emphasizing their relevance to

our study population. Thank you again for prompting this refinement, which strengthens the methodological rigor of our work.

7.Method, data analysis - Suggest that the authors explore preference heterogeneity, through inclusion of interaction terms and/or mixed logit modelling. It is possible that different patient subgroups have different preferences for health state attributes and it would be informative to understand the level of variation within the sample.

Reply： Thank you profoundly for this invaluable suggestion to explore preference heterogeneity—your insight not only aligns perfectly with our own research goals but even resonates with the original intent behind how we designed our study sample. From the outset, we recognized that capturing the full complexity of health preferences would require examining both inter-group differences (between cancer patients and the general population, which is the focus of this manuscript) and intra-group variations within patient subgroups. In fact, exploring preference heterogeneity among cancer patients was explicitly the second core objective in designing our research sample. Your recommendation to delve into this heterogeneity feels like a serendipitous convergence of ideas, and we are genuinely excited by how it reinforces the depth of inquiry we aimed to achieve from the start.

To clarify, the current manuscript specifically focuses on comparing QLU-C10D preferences between Chinese cancer patients and the general population—a targeted analysis of inter-group differences. However, we fully agree that exploring subgroup variations within patient cohorts is equally critical, and this aligns seamlessly with your suggestion.

Notably, the companion study dedicated to this topic has already been completed, and it has been developed in parallel with the current manuscript. Using mixed logit modeling and interaction terms, it dissects preference heterogeneity among cancer patients, with the tentative title "Exploring Preference Heterogeneity in QLU-C10D Health State Valuations: Variations Across Cancer Types and Treatment Modalities." This study investigates how preferences for health attributes differ across subgroups (e.g., by cancer type, age, and treatment modality).

Preliminary findings include, for example, that younger breast cancer patients prioritize "emotional function" more than older patients, while elderly lung cancer patients value "physical function" higher. Patients on immunotherapy also place greater weight on "skin symptoms" compared to those receiving chemotherapy—likely due to distinct adverse effect profiles. These insights complement the current manuscript, and we plan to reference them in future discussions to provide a more comprehensive picture.

Thank you again for this astute suggestion, which not only enhances our work but also validates the foundational vision we had for this research. We truly look forward to the possibility of you encountering this companion study on heterogeneity in the future—it would mean a great deal to us to have your perspective on it.

8. Method, data analysis - The authors state that "Non-monotonic coefficients were adjusted by constraining adjacent levels, as per standard practice." The choice to constrain attribute levels should be supported by the results of likelihood-ratio testing indicating that the part-worth utility values for the levels are not significantly different from one another. Suggest reporting on these hypothesis test results.

Reply： Thank you profoundly for this insightful suggestion, which has given us a new perspective we hadn't previously considered. Your point about validating the constraint of adjacent levels through likelihood-ratio testing is crucial for enhancing the rigor of our methodology, and we truly appreciate you highlighting this.

Following your guidance, we conducted supplementary analyses in Stata, specifically performing a likelihood-ratio test on the non-monotonicity observed between role functioning levels 3 and 4—with level 3 as the reference. The results showed no statistically significant difference between these two levels ($p = 0.21$), confirming that participants' preferences for them are not distinct. This validates the rationale for imposing monotonicity constraints to resolve the inconsistency.

We have accordingly added the following to the manuscript: "A likelihood-ratio test confirmed no significant difference between these levels, so monotonicity was imposed and the model re-estimated."

What's more, this approach to validating constraints through statistical testing is likely to inform many of our future studies, as it sets a stronger standard for methodological rigor. Thank you again for this invaluable input, which has significantly strengthened our work.

9. Method, data analysis - Suggest including how the utility value for any health state was calculated based on the estimated utility decrements for transparency.

Reply： Thank you for suggesting that we clarify how utility values for health states are calculated based on estimated utility decrements, to enhance transparency. Your suggestion has made us realize that this detail is indeed crucial for ensuring the methodology is understandable and reproducible.

We have added a specific example in the last part of "3.3 Utility estimates" to illustrate this process:

" For example, for QLQ-C30 responses matching the QLU-C10D health state "3233412432," the utility score is calculated as: 1 - 0.275 - 0.052 - 0.035 - 0.057 - 0.125 - 0 - 0.046 - 0.083 - 0.075 - 0.030 = 0.222. The worst possible health state (PITS state 4444444444) had a utility value of −0.092, considerably lower than the Chinese general population PITS value (0.083)[45]."

This example, provided after establishing the value set, aims to clearly demonstrate the calculation logic. We appreciate your guidance in improving the clarity of our methodology.

10. Discussion - This study compares results from non-linear analyses across two populations. Suggest speaking to scale heterogeneity in this setting (DOI: 10.1007/s40271-017-0282-4) and comparisons when models upon which utility decrements were based included different underlying parameters (given differences in non-monotonicity observed across populations varied and required different constraints).

Reply： Thank you for highlighting the need to address scale heterogeneity and model parameter differences in our cross-population comparison. Your reference to Rose et al. (2017) has been incredibly illuminating—frankly, your consecutive suggestions have opened up fascinating new angles we hadn't fully explored, making the study feel richer and more layered with each insight. It's exciting to see how much more there is to unpack!

As suggested, we've added a concise note in the discussion to acknowledge these factors:

" It is important to note that cross-population differences could also reflect scale heterogeneity (i.e., variations in choice consistency) [70], as well as differences in model constraints applied to address non-monotonicities between groups. These factors may influence utility estimates beyond true preference differences, warranting cautious interpretation."

This revision frames potential influences on our findings without overcomplicating the discussion, as you recommended. Thank you again for these eye-opening inputs—they've not only strengthened our interpretation but also sparked new curiosity about deeper layers we can explore.

11. Discussion - This sentence should include a reference for global incidence numbers "This study developed a QLU-C10D value set based on the health preferences of cancer patients in China, the country with the highest cancer incidence globally, and compared them with the QLU-C10D value set of the general public in China derived."

Reply： Thank you for highlighting the need to include a reference for the global incidence figures in the discussion. We have supplemented the relevant sentence with the cited reference as follows:

"This study developed a QLU-C10D value set based on the health preferences of cancer patients in China, the country with the highest cancer incidence globally [Han B, Zheng R, Zeng H, Wang S, Sun K, Chen R, Li L, Wei W, He J: Cancer incidence and mortality in China, 2022. J Natl Cancer Cent 2024, 4(1):47-53], and compared it with the QLU-C10D value set derived from the general public in China..."

This addition ensures that the claim regarding global cancer incidence is backed by authoritative data, strengthening the validity of the discussion. We appreciate your attention to this important detail.

12. Discussion - Suggest contextualizing findings within broader PRO valuation literature (which results were/were not comparable to patient vs population comparisons from other PRO valuations, e.g, EQ-5D, SF-6D).

Reply： Thank you sincerely for this thoughtful suggestion. We fully recognize the importance of contextualizing our findings within the broader PRO valuation literature, including comparisons across instruments like EQ-5D and SF-6D, and we greatly appreciate your attention to this critical aspect of our work.

As elaborated in our previous publication “Chinese utility weights for the EORTC cancer-specific utility instrument QLU-C10D”, we made detailed comparisons between the QLU-C10D general population value set and generic MAUIs, specifically the EQ-5D-5L and SF-6Dv2. In terms of similarities, we found that the greatest utility decrement occurs in Physical functioning and Pain in both cancer-specific (QLU-C10D) and generic MAUIs. This is not unexpected considering that all value sets are grounded in the preferences of the general public in China.

However, differences also emerged. The utility value ranges varied significantly. The Chinese EQ-5D-5L and SF-6Dv2 value ranges are [-0.391,1] and [-0.277,1] respectively, which are wider than that of QLU-C10D. These disparities can be ascribed to the distinct health state classification systems and valuation methods. The two Chinese generic value sets were developed using a TTO (Time Trade-Off) technique, which requires weighing the duration of health status, while our study on QLU-C10D adopted a DCE (Discrete Choice Experiment) design. As is well-known, valuation methods have a substantial impact on the resultant utility values [56].

In the current manuscript, we have further strengthened the contextualization by delving into nuanced preference differences, particularly around emotional functioning and the lived experience of patients. For example, we highlight that in the Chinese general population value set, emotional functioning exhibits the smallest utility decrement [45]. This discrepancy may be attributed to China's unique disease culture and traditional Chinese medicine belief that "emotions can cause cancer, and emotions can heal cancer" [68]. Additionally, we emphasize that utility decrements from the cancer patient value set exceed those from the general population across most dimensions, particularly symptom-related ones. The consequence is that the cancer patient health state values are generally lower than those of the general population, with the maximum difference at the pit state (-0.092 versus 0.083). The most likely explanation for these preference differences is that patients have a nuanced understanding of health impairments through their lived experience, whereas the general population tends to underestimate the impact of symptoms due to a lack of firsthand experience [9, 69].

Our focus in the current manuscript remains on comparisons within the QLU-C10D framework (patient vs. general population value sets) and cross-country comparisons of cancer-specific value sets. This approach was intended to avoid redundancy with our prior work, allowing each study to enhance its unique contribution. The previous study placed QLU-C10D in the context of generic instruments, while the current one zeroes in on patient-specific preferences, a crucial yet under-explored dimension in oncology PRO valuation.

To ensure readers can relate our findings to the broader literature, we have clearly referenced our prior publication in the Discussion section, guiding readers to that work for comprehensive cross-instrument comparisons. This way, the current manuscript can maintain its focus on cancer patient preferences for QLU-C10D while still providing access to the broader context of PRO valuation.

We hope this approach balances the need for contextual richness and the aim of avoiding redundancy, and we are open to further adjustments if necessary. Thank you again for your meticulous consideration of our work—your insight has helped us strengthen the depth and relevance of our analysis.

13.Discussion - Additional study limitations should be transparently reported (e.g., limited recruitment to two hospitals, hypothetical bias for stated preference methods).

Reply： Thank you for your valuable suggestion to ensure all relevant study limitations are transparently reported. We fully agree that this transparency is critical for readers to contextualize our findings and assess the robustness of our work.

As requested, we have thoroughly addressed the limitations you highlighted, alongside other key constraints, in the discussion section. Our complete limitations statement is as follows:

"However, several limitations should be acknowledged. Firstly, patient recruitment was limited to two tertiary public hospitals in Harbin, China. While this approach ensured data quality and feasibility, it may limit the generalizability of our findings, as the recruited sample might not fully represent the broader cancer patient population across China in terms of socioeconomic, regional, and clinical diversity. Secondly, voluntary participation may have introduced sample bias, potentially favoring individuals with milder conditions or better treatment outcomes, while close interviewer involvement could have induced some social desirability bias. Thirdly, we did not apply education weighting (as used in the general population value set) due to inherent differences in education distributions between cancer patients and the general population; this omission may weaken the stability of extrapolating our findings for cross-group comparisons. Finally, as with all stated preference methods, our use of discrete choice experiments may be subject to hypothetical bias, where participants' stated choices may not perfectly reflect real-world decision-making."

This includes explicit mention of the two points you emphasized: the limited recruitment to two hospitals (first limitation) and the hypothetical bias inherent to stated preference methods (fourth limitation). By presenting these alongside other relevant constraints—such as voluntary participation bias and the absence of education weighting—we aim to provide a comprehensive and transparent account of the study's limitations.

Thank you again for reinforcing the importance of this thoroughness, which strengthens the integrity of our research.

Quality of Life Utility-Core 10 Dimensions (QLU-C10D) scores based on cancer patients' health preferences are more sensitive than those based on general public's health preferences

Yiyin Cao[1]*, Haofei Li[1]*, Xu Jin[1], Madeleine T. King[2], Richard Norman[3], Ling Jie Cheng[4,5,6], Hongjuan Yu[7#], Weidong Huang[1#], Nan Luo[6]

[1] School of Health Management, Harbin Medical University, Harbin, China

[2] School of Psychology, University of Sydney, Sydney, Australia

[3] School of Public Health, Curtin University, Perth, Australia

[4] Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom

[5] Yong Loo Lin School of Medicine, National University of Singapore, Singapore

[6] Saw Swee Hock School of Public Health, National University of Singapore, Singapore

[7] Department of Hematology, The First Affiliated Hospital of Harbin Medical University, Harbin, China

*These authors have made equal contributions and shared first authorship.

[#]Correspondence:

Weidong Huang: huangweidong@hrbmu.edu.cn

Hongjuan Yu:yuhongjuan2008@163.com

Weidong Huang is the leading corresponding author

Declarations

Conflict of interest: The authors declare that they have no conflicts of interest.

Funding/Support: This work was supported by the National Science Foundation of China (Grant No. 72274045, 71974048)

Ethics approval statement: The protocol of this cohort study received approval from the Ethics Committee of Harbin Medical University (HMUIRB2023005) and was conducted in accordance with the principles outlined in the Declaration of Helsinki.

Authors' contributions: Conceptualization: Weidong Huang, Madeleine T. King, Hongjuan Yu, Nan Luo; Methodology: Yiyin Cao, Haofei Li, Richard Norman, Ling Jie Cheng; Formal analysis and investigation: Yiyin Cao, Haofei Li, Xu Jin, Hongjuan Yu; Writing - original draft preparation: Yiyin Cao, Weidong Huang, Madeleine T. King, Nan Luo; Writing - review and editing: Madeleine T. King, Richard Norman, Hongjuan Yu, Weidong Huang, Nan Luo; Funding acquisition: Weidong Huang;Supervision: Nan Luo, Madeleine T. King

Quality of Life Utility-Core 10 Dimensions (QLU-C10D) scores based on cancer patients' health preferences are more sensitive than those based on general public's health preferences

Abstract

Objective: This study aimed to develop a value set for the European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Utility-Core 10 Dimensions (QLU-C10D) based on the preferences of Chinese cancer patients, and to compare results with the Chinese general population value set.

Methods: Patients with mixed cancer diagnoses were recruited from two hospitals in China between September and November 2024. Patients completed the EORTC Core Quality of Life Questionnaire (QLQ-C30) followed by a discrete choice experiment (DCE) valuation survey consisting of 16 choice sets through face-to-face interviews. A conditional logit model was applied to analyze patient preferences. The characteristics and sensitivity of the derived QLU-C10D value set were compared with those of the Chinese general population QLU-C10D value set. Sensitivity was evaluated by ability to discriminate cancer stage and performance status, and effect sizes (ES) were estimated.

Results: Data from 659 cancer patients were analyzed. The cancer patients' value set showed a broader utility range (-0.092 to 1.0) compared to the general population's value set (0.083 to 1). The mean utility score for the cancer patient sample (0.656) was significantly lower than that for the general population (0.738, $P < 0.01$). Both value sets demonstrated good sensitivity, but the cancer patients' value set was more effective in distinguishing clinical subgroup differences (ES = 2.76–4.46 vs. ES = 2.47–3.48).

Conclusion: Differences in value sets between cancer patients and the general population underscore the importance of more comprehensively integrating patient perspectives into existing evaluation frameworks.

Keywords: Value Set, Cancer Patients, General Population, QLU-C10D, China

1 Introduction

Cost-utility analysis (CUA) is a key method for reimbursement decisions in healthcare, typically using preference-based health-related quality of life (HRQoL) utility values to derive quality-adjusted life years (QALYs). The QALY combines morbidity and mortality into a single metric, weighed by preferences that reflect the relative desirability of health states, serving as the basis for evaluating the impact of health interventions [1]. Country-specific value sets, reflecting societal preferences and collective benefits, are considered ideal for healthcare decision-making [2, 3]. This practice, rooted in the principle of maximizing societal value, has been widely endorsed by economic evaluation guidelines across multiple countries [4].

However, in the context of patient-centered decision-making, the debate over prioritizing general population versus patient-derived preferences remains unresolved [4-6]. While general population preferences capture societal benefits, patient preferences represent the lived experiences associated with specific health states. In Germany, the HTA framework led by the Institute for Quality and Efficiency in Health Care focuses on qualitative assessments of "added clinical benefit" from the patient perspective, such as survival gains and symptom improvement. It directly incorporates patient-reported outcomes (PROs) to evaluate therapies against comparators, ensuring patient voices are central to treatment benefit evaluations[7]. Similarly, Sweden's Dental and Pharmaceutical Benefits Agency emphasizes ethical principles like "human value and need" and "solidarity" in its health economic evaluations, which prioritize patient experiences. These evaluations integrate condition-specific PROs to align resource allocation with patients' actual needs [8]. By incorporating patient preferences, HTAs can better align resource allocation with actual patient needs, improve perceived health benefits and satisfaction, and enhance personalized healthcare [9-11]. Therefore, integrating value sets representing preferences informed by the lived experience alongside societal perspectives offers a more comprehensive framework for clinical and policy decision-making.

Significant preference differences highlight the critical importance of choosing between patient and general population preferences. A meta-analysis conducted in 2006 indicated that minimal or no significant differences between utility values derived from these groups [5]. However, recent studies increasingly reveal notable discrepancies between these groups, encouraging the development of patient-specific value sets. For example, some studies suggest that utility values derived from patients tend to be systematically higher [12, 13], while others report the opposite trend [14-16]. Moreover, the direction and magnitude of these differences appear to be influenced by factors such as disease type and national context [17]. These findings reflect the unique experiences, recovery expectations, and treatment perceptions of patients, highlighting the limitations of universally applying general population-based value sets in healthcare decision-making. This calls for further exploration into the role of patient preferences in health evaluations and their divergence from general preferences.

In recent years, the European Organization for Research and Treatment of Cancer (EORTC), in collaboration with the Multi-Attribute Utility in Cancer (MAUCa) Consortium, developed the QLU-C10D [18], a cancer-specific preference-based measure (PBM) derived from the widely used Quality of Life Questionnaire - Core 30 (QLQ-C30) [19, 20]. In China, value sets based on general population preferences have been developed for several generic PBMs, including the EQ-5D [21-24], EQ-5D-Y [25], and SF-6D [26]. While these PBMs are widely applied in HTA, they are not cancer-specific and may not fully capture the symptom burden and HRQoL changes experienced by cancer patients. The QLU-C10D, by contrast, has shown high sensitivity in detecting subtle health changes and distinguishing between

health states among cancer patients [27-32]. Value sets for the QLU-C10D, developed from general population preferences, are now available in 16 countries [33-45], including China, facilitating its application in HTA [46-49]. However, patient-derived QLU-C10D value sets have received increasing attention only recently. Gamper et al. were the first to apply the QLU-C10D valuation survey among cancer patients in Austria, confirming its feasibility and laying the groundwork for patient-specific value sets [10]. Gandhi et al. extended this work by comparing utilities derived from general population and cancer patient preferences in Singapore [50]. However, the impact of such differences on CUAs remains underexplored, warranting further research to advance personalized healthcare and refine economic evaluation methods.

This study aimed to develop QLU-C10D utility weights for cancer patients in China, the country with the highest global cancer incidence [51], to inform clinical decision-making. Additionally, it aimed to explore differences in Chinese QLU-C10D value sets based on health preferences from cancer patients versus the general population, and consider how these differences may impact CUA.

2 Methods

The research design and analytical approach in this study follow the established QLU-C10D valuation methods used in previous studies, which are based on general population preferences from multiple countries [33-45]. These methods generally adhere to the standardized EORTC protocol for value set endorsement, with two adjustments for the patient context: recruitment and administration of the valuation study. Previous studies have demonstrated the feasibility of conducting QLU-C10D valuations among cancer patients [10].

2.1 Sample recruitment and management

From September to November 2024, the study recruited hospitalized cancer patients from two Grade A tertiary public hospitals in Harbin, the capital of Heilongjiang Province, including a cancer specialty hospital. The aim was to ensure a diverse sample in terms of age, residence, diagnosis, and treatment modalities. Participants were selected by hospital doctors or nurses based on the following criteria: (1) confirmed clinical diagnosis of cancer; (2) anticipated survival of at least one year for patient safety and ethical compliance; (3) age ≥18 years; and (4) proficiency in reading and communicating in Chinese.

Eligible patients provided written informed consent and participated in face-to-face interviews conducted by trained interviewers in hospital wards. First, as a familiarization exercise, the patient self-completed the QLQ-C30, albeit with the items read to them by an interviewer. The interviewer then explained the discrete choice experiment (DCE) task, using example questions to ensure comprehension. Responses were recorded electronically, with interviewers providing clarifications when needed. Standardized training was implemented, and a protocol was strictly followed to minimize interviewer bias and ensure data quality.

2.2 The QLU-C10D health state classification

The scenarios in the DCE focused on the QLU-C10D dimensions, derived from the QLQ-C30 by the MAUCa Consortium (2010–2016) using item response theory, qualitative interviews with cancer patients, expert opinion of oncologists and HRQOL researchers [18]. The QLU-C10D includes 13 items across ten health dimensions—physical, role, social, and emotional functioning; pain; fatigue; sleep disorders; appetite; nausea; and bowel problems. Each dimension has four severity levels (reflecting the source QLQ-C30 items levels), defining 1,048,576 unique health states [18]. Table A (supplementary material) details the classification system.

2.3 Valuation survey and DCE design

The valuation survey followed the original methods designed by the MAUCa Consortium and EORTC QOL Group [33], modified only in recruitment and implementation methods to enable cancer patient involvement. The DCE design and data management were supported by Survey Engine, and complied with the International Code on Market, Opinion and Social Research and Data Analytics [52].

Participants completed a survey comprising four sections: (1) self-reported health scales (with assistance of interviewers for reading and entering responses only), including QLQ-C30 [19], Kessler Psychological Distress Scale (Kessler-10) [53] for mental health, EQ-5D-5L [54], and SF-6Dv2 [55], all using officially endorsed Chinese versions; (2) DCE tasks with 16 choice sets; (3) feedback on DCE tasks, covering task difficulty, clarity, and choice strategies; and (4) sociodemographic questions, such as gender, age, and residence. Clinical data, including diagnosis, cancer staging, treatment modalities, and ECOG performance status, were obtained from medical records [56].

This study employed the DCE-based QLU-C10Dvaluation protocol described by King et al. [33], widely used thereafter in valuation studies across numerous countries [33-45]. Patients completed 16 choice sets, selected from 960 combinations via a D-efficiency procedure (a key method under optimal design theory) that prioritizes maximizing the efficiency of main effect parameter estimates. Interaction effects were not included to balance survey feasibility and statistical precision, consistent with common practices [33-45]. To control biases, health state A/B assignment and the order of 11 attributes across 10 QLU-C10D domains (the physical functioning dimension was split into 2 attributes, one for 'long walk' and one for 'short walk' (see Table A from supplementary material) were randomized per respondent; a balanced incomplete block design (BIBD) defined four differing dimensions per set (others held equal) to minimize burden. Each set presented two hypothetical QLU-C10D health states, differing in four of ten HRQoL dimensions and survival times (1, 2, 5, or 10 years), with differences highlighted for clarity. Examples are shown in Figures A and B (Chinese version) in the supplementary material. Details on QLU-C10D methodology, including formats, attribute ordering, and reliability, are provided in related studies [10, 57-59].

2.4 Data quality control

First, face-to-face interviews were conducted to optimize survey comprehension and interactivity: trained interviewers clarified DCE tasks, offered assistance as required, and halted surveys exclusively under specific circumstances—including inability to understand the survey interface, persistent failure to grasp task requirements (despite repeated explanations), or voluntary requests to withdraw (e.g., due to fatigue). These situations were deemed to directly compromise data validity. A total of 7 surveys (1.06% of the initial sample) were terminated, resulting in a final sample of 659 participants. The primary reasons for termination were fatigue (n=5) and persistent comprehension difficulties (n=2). Notably, none of these terminations were due to uncertainty or difficulty in making choices during the choice tasks. Secondly, patients who provided the most negative responses to all four DCE feedback questions ("Survey is more difficult," "Health states very unclear," "Choices very difficult," and "No choice strategy") were excluded as part of a robustness check to ensure task comprehension. Lastly, to assess the impact of survey duration on data quality, respondents were divided into deciles by completion time. Pseudo-R^2 values and significant coefficients from conditional logit analysis were plotted to examine trends.

2.5 Data analysis

Descriptive statistics summarized the characteristics of the cancer patient sample; specifically,

frequency and percentage for categorical variables, mean and standard deviation for continuous variables.

Conditional logistic regression was employed to model utility and calculate utility weights by interacting QLU-C10D dimensions with survival time. Utility values were anchored at 0 ("dead") and 1 ("full health") to maintain ratio-scale properties[60, 61],yet remained unconstrained during initial estimation to capture states "worse than deaths". The utility of each option was calculated as follows:

$$U_{isj} = \alpha \mathrm{TIME}_{isj} + \beta X'_{isj} \mathrm{TIME}_{isj} + \varepsilon_{isj} \quad (1)$$

where TIME_{isj} represents the survival time presented in DCE option j, and X'_{isj} was a set of dummy variables corresponding to the levels of the health state in DCE option j. The error term, εisj, was assumed to follow a Gumbel distribution. Conditional logistic regression estimated parameters α (utility per life year) and β (utility weights for each health state level). Utility decrements were calculated as the ratio of β to α, reflecting the HRQoL-life expectancy trade-off. Non-monotonic coefficients were adjusted by constraining adjacent levels, as per standard practice[62].

Line charts were used to compare utility decrements across QLU-C10D dimensions and levels between the cancer patient value set versus the Chinese general population value set. The general population value set served as the baseline to illustrate the percentage changes in utility decrements for the cancer patient value set, visualized using bar charts. Additionally, 1,000 hypothetical health states were generated by randomly combining the 4 levels across the 10 dimensions of the QLU-C10D descriptive system. This sampling approach ensured proportional representation of each dimension's levels, with states spanning the full severity spectrum—from mild (e.g., most dimensions at levels 1–2) to severe (e.g., majority at levels 3–4)—and explicitly including the best-possible (all dimensions at level 1) and worst-possible (all dimensions at level 4) health states. This comprehensive coverage of potential health states was designed to ensure robust utility comparisons between the two value sets.

Box plots displayed utility scores derived from both value sets, highlighting ranges and means. Agreement was assessed using intraclass correlation coefficients (ICC; high > 0.75) and Bland-Altman plots [63]. Sensitivity analysis stratified patients by cancer stage (I–IV) and ECOG performance status (Grades 0–4). Mean utility scores for each subgroup were compared using one-way ANOVA. "Extreme subgroups" here refer to the two ends of the stratified variables: for cancer stage, these were Stage I (earliest) and Stage IV (most advanced); for ECOG performance status, they were Grade 0 (fully active) and Grade 4 (completely disabled). Effect sizes (ES) quantified the clinical significance of differences between these extreme subgroups, calculated as the mean score difference divided by their pooled standard deviation [64]. Incremental utility was defined as the absolute difference in mean utility values between these extreme subgroups, reflecting the utility change associated with moving from the most severe to the least severe end of the stratified variable. We hypothesized higher utility values for earlier cancer stages or lower ECOG grades[29, 65-67].

3 Results

3.1 Sample characteristics

Of the 727 participants, 68 were excluded due to missing medical records or incomplete disease information, leaving 659 patients with various cancer types. The most common cancers were lung, breast stomach, and thyroid. Cancers with a prevalence of less than 3% were grouped into the 'other' category, which accounted for 7% of the sample. The sample included 347 males (53%) and 312 females (47%), with 48% aged 60 or older, 61% residing in urban areas, 69% having a junior high education or below, and 91% married. Detailed socio-demographic characteristics are presented in Table 1.

Table 1 Socio-demographic characteristics of the study samples

	Cancer patient sample (n=659)	
	N	%
Gender[a]		
Male	347	52.7
Female	312	47.3
Age group		
18-39	53	8.0
40-59	287	43.6
≥60	319	48.4
Residence		
Urban	402	61.0
Rural	257	39.0
Education level		
Junior high school or below	454	68.9
Senior high school/technical school	118	17.9
Post-secondary	87	13.2
Marital status		
Single	21	3.2
Married	602	91.4
Widowed/divorced/separated	36	5.5
Cancer type		
Lung cancer	145	22.0
Breast cancer	108	16.4
Stomach cancer	87	13.2
Thyroid cancer	67	10.2
Lymphoma	53	8.0
Leukemia	46	7.0
Liver cancer	44	6.7
Colorectal cancer	33	5.0
Prostate cancer	28	4.2
Others	48	7.3

[a] No significant difference was found between the distribution in this study and the latest incidence rates reported by the National Cancer Center and the International Agency for Research on Cancer (IARC). (/www.iarc.who.int/wp-content/uploads/2024/02/pr345_E.pdf) ($\chi^2 = 0.343$, $p = 0.558$)

3.2 Data quality control

Most patients found the DCE survey difficulty comparable to other surveys (79%), with few citing unclear health state descriptions (2%) or difficulty choosing between health states (14%, "very difficult" 0.2%). A majority (69%) focused on the "highlighted aspects" of the tasks. Detailed feedback results are shown in Figure C (supplementary material). No participants were excluded for consistently negative responses.

The mean survey completion time was 29.95 minutes (range: 5.39–80.12). Completion times decreased with task familiarity (Figure D in the supplementary material). The median decile produced the highest number significant coefficients (25/31). The fastest decile produced the fewest significant coefficients (6/31) and yet achieved comparable or higher model fit (pseudo-R^2 value 0.442) than other deciles (Figure E in the supplementary material); for further detail and discussion see the supplementary material. These results confirm the dataset's high quality and reliability.

3.3 Utility estimates

As shown in the "Unconstrained" column of Table 2, incremental moves to worse levels within each dimension generally led to significantly greater coefficients. However, non-monotonicities were observed in role functioning levels 3 and 4, where level 4 was preferred over level 3. A likelihood-ratio

test confirmed no significant difference between these levels, so monotonicity was imposed and the model re-estimated. Utility estimates for the cancer patient sample after correction show similar utility weights in both unconstrained and monotonicity-corrected model.

Monotonicity-corrected results were used to develop the QLU-C10D value set for cancer patients in China. Utility scores are derived by mapping QLQ-C30 responses to the corresponding QLU-C10D dimensions and levels (Table A in the supplementary material). For example, for QLQ-C30 responses matching the QLU-C10D health state "3233412432," the utility score is calculated as: 1 - 0.275 - 0.052 - 0.035 - 0.057 - 0.125 - 0 - 0.046 - 0.083 - 0.075 - 0.030 = 0.222. The worst possible health state (PITS state 4444444444) had a utility value of −0.092, considerably lower than the Chinese general population PITS value (0.083)[45].

Table 2 QLU-C10D model coefficients and utility decrements for the cancer patient sample

		Unconstrained			Corrected for monotonicity		
Parameter	**Level[a]**	**Coefficient (α, β)**	**SE**	**Utility decrement (β/ α)**	**Coefficient (α, β)**	**SE**	**Utility decrement (β/ α)**
Time coefficient (α)	(linear)	0.885**	0.059		0.882 **	0.059	
Attributes (β)							
Physical Functioning	2	-0.145 **	0.026	-0.164	-0.145 **	0.027	-0.165
Physical Functioning	3	-0.242 **	0.032	-0.273	-0.242 **	0.032	-0.275
Physical Functioning	4	-0.300 **	0.032	-0.339	-0.301 **	0.032	-0.341
Role Functioning	2	-0.048 **	0.013	-0.054	-0.046 **	0.013	-0.052
Role Functioning	3	-0.098 **	0.014	-0.111	-0.091 **	0.013	-0.103
Role Functioning	4	-0.087 **	0.013	-0.098	-0.091**	0.013	-0.103
Social Functioning	2	-0.014	0.012	-0.016	-0.015	0.012	-0.017
Social Functioning	3	-0.031 *	0.013	-0.035	-0.031 *	0.013	-0.035
Social Functioning	4	-0.057 **	0.012	-0.064	-0.056 **	0.012	-0.064
Emotional Functioning	2	-0.024 *	0.012	-0.027	-0.024 *	0.012	-0.027
Emotional Functioning	3	-0.050 **	0.013	-0.057	-0.050 **	0.013	-0.057
Emotional Functioning	4	-0.096 **	0.012	-0.108	-0.096 **	0.012	-0.109
Pain	2	-0.034 **	0.013	-0.039	-0.034 **	0.012	-0.039
Pain	3	-0.094 **	0.013	-0.106	-0.093 **	0.013	-0.106
Pain	4	-0.111 **	0.012	-0.125	-0.110 **	0.012	-0.125
Fatigue	2	-0.035 **	0.011	-0.040	-0.035 **	0.011	-0.040
Fatigue	3	-0.055 **	0.013	-0.062	-0.054 **	0.012	-0.061
Fatigue	4	-0.061 **	0.011	-0.068	-0.060 **	0.011	-0.068
Sleep disorders	2	-0.041 **	0.011	-0.046	-0.041 **	0.011	-0.046
Sleep disorders	3	-0.051 **	0.012	-0.058	-0.051 **	0.012	-0.058
Sleep disorders	4	-0.057 **	0.010	-0.065	-0.057 **	0.010	-0.064
Lack of appetite	2	-0.042 **	0.011	-0.047	-0.041 **	0.011	-0.047
Lack of appetite	3	-0.055 **	0.013	-0.062	-0.054 **	0.013	-0.061
Lack of appetite	4	-0.074 **	0.012	-0.083	-0.073 **	0.012	-0.083
Nausea	2	-0.039 **	0.012	-0.044	-0.040 **	0.012	-0.045
Nausea	3	-0.067 **	0.013	-0.075	-0.066 **	0.013	-0.075
Nausea	4	-0.072 **	0.012	-0.081	-0.071 **	0.012	-0.080
Bowel problems	2	-0.027 *	0.012	-0.030	-0.026 *	0.012	-0.030
Bowel problems	3	-0.035 **	0.012	-0.039	-0.034 **	0.012	-0.039
Bowel problems	4	-0.049 **	0.011	-0.055	-0.049 **	0.011	-0.055

* p<0.05; ** p<0.01;

[a] Level 2= a little; Level 3 = quite a bit; Level 4 = very much; SE= Standard

3.4 Comparison of value sets from cancer patients versus the general population

Utility decrements increased monotonically with higher severity levels in most dimensions in both value sets. The cancer patient value set had more dimensions with monotonic decrements than the general population value set, requiring fewer corrections. Functional dimensions (e.g., physical, emotional) exhibited larger utility decrements than symptom dimensions in both value sets. Physical functioning had by far the biggest utility decrements in both value-sets, and pain ranked second in both. Emotional functioning ranked third in the cancer patient value set, while role functioning ranked third in the general population value set (Figure 1). For the symptoms, the patient data consistently yielded larger decrements than the general population data, except for pain. The standard errors for cancer patients' utility decrement—visualized as error bars in Figure 1—occasionally exceed these observed differences, with such error bars reflecting statistical uncertainty.

Figure 2 compares utility decrements across severity levels from the cancer patient value set and the general population value set (with the baseline set at 0), highlighting absolute differences. Cancer patients showed significantly larger decrements across all dimensions, except for certain levels of role and social functioning and sleep disorders, and all levels of pain. The largest differences were observed at level 2, particularly in symptom dimensions like fatigue, lack of appetite, and bowel problems.

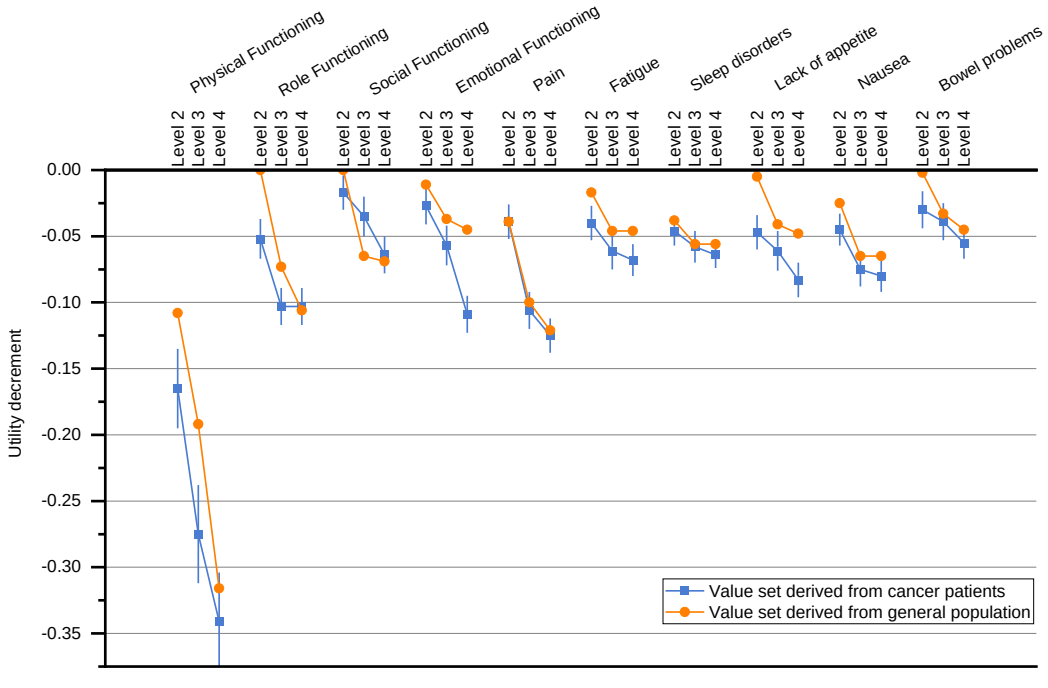

Fig. 1 QLU-C10D value sets for the cancer patient sample and the general population

Note: The value set derived from the general population is based on data from a published study[45].

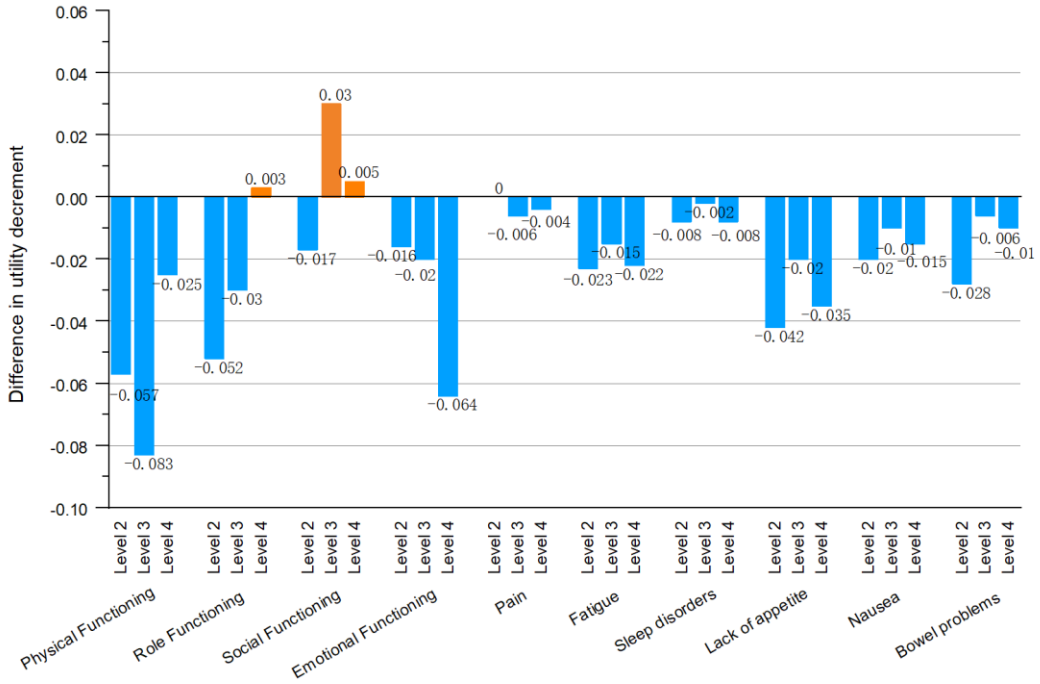

Fig. 2 Difference in value sets between cancer patients and the general population

Note: The baseline (0) represents the value set derived from the general population. The value set derived from the general population is based on data from a published study[45].

Figure 3 shows utility values for 1,000 randomly selected QLU-C10D health states, estimated using value sets from cancer patients and the general population, and ordered according to the values from the general population. Utility values from the cancer patient value set were consistently lower, especially for better health states.

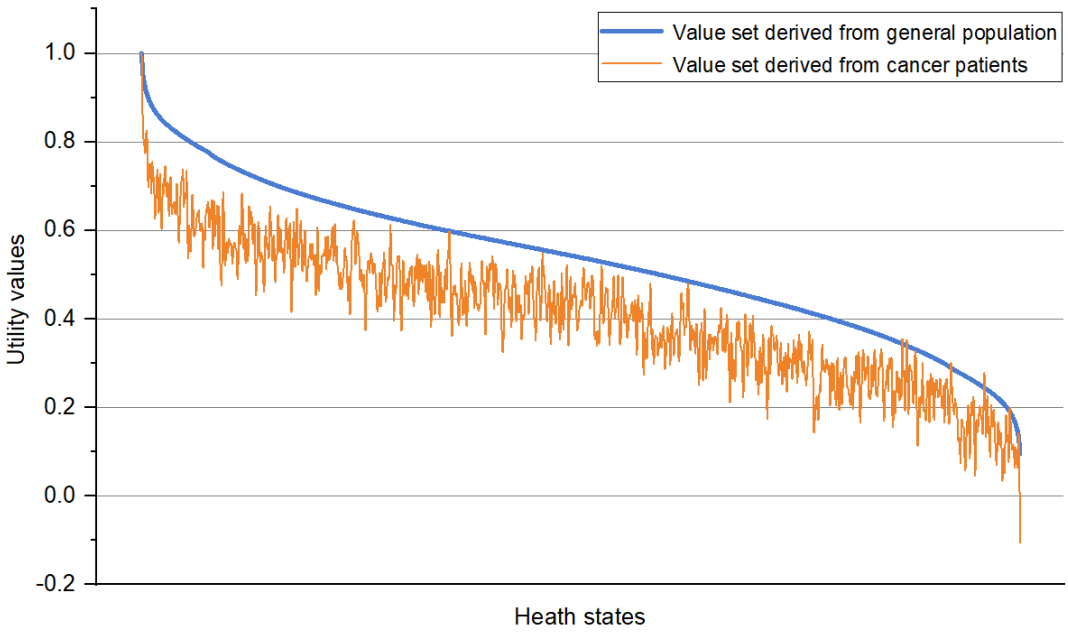

Fig. 3 QLU-C10D health states values estimated using value sets derived from Chinese cancer patients and the Chinese general population (based on data from a published study[45]), ordered according to the

values from the general population.

Figure F in the supplementary material presents utility value distributions for cancer patients using both patient and general population value sets. The cancer patient value set had a broader range (-0.035 to 1.0) and lower mean utility score (0.656) compared to the general population value set (0.091 to 1.0, mean score: 0.783), with a significant difference (P<0.01). High agreement was observed (ICC = 0.87). The Bland-Altman plot (Figure G in the supplementary material) showed proportional bias, with 95% limits of agreement ranging from -0.30 to 0.05.

3.5 Sensitivity comparison between value sets derived from cancer patients and the general population

Utility values derived from the general population and cancer patient value sets differed significantly (p<0.01) across groups stratified by cancer stage and ECOG performance status (Table 3). ES estimates comparing extreme subgroups for the cancer patient value set were larger than those for the general population for cancer stage and ECOG performance status, with larger incremental utilities.

Table 3 Sensitivity for value sets derived from Chinese cancer patients and general population

Preferences	Cancer stage	N	Mean (SD)	*p*-value	ES	Incremental utilities
Cancer patients	I	195	0.84±0.14			
	II	252	0.68±0.20			
	III	137	0.40±0.12			
	IV	25	0.26±0.23	<0.01	4.46	0.49
General population	I	195	0.90±0.09			
	II	252	0.81±0.13			
	III	137	0.60±0.13			
	IV	25	0.47±0.16	<0.01	3.48	0.33
Preferences	**ECOG**	**N**	**Mean (SD)**	***p*-value**	**ES**	**Incremental utilities**
Cancer patients	Grade 0	154	0.87±0.17			
	Grade 1	361	0.64±0.21			
	Grade 2	111	0.49±0.21			
	Grade 3	28	0.42±0.23			
	Grade 4	5	0.39±0.18	<0.01	2.76	0.48
General population	Grade 0	154	0.92±0.12			
	Grade 1	361	0.78±0.14			
	Grade 2	111	0.65±0.17			
	Grade 3	28	0.56±0.22			
	Grade 4	5	0.55±0.19	<0.01	2.47	0.37

ES: Effect size; SD: Standard deviations

Note: The analyses of cancer stage excluded leukemia cases (N=609), whereas all other analyses were conducted on the full sample (N=659).

4 Discussion

This study developed a QLU-C10D value set based on the health preferences of cancer patients in China, the country with the highest cancer incidence globally [51], and compared them with the QLU-C10D value set of the general public in China derived. This enabled us to examine differences between cancer patient versus general population value sets within the same geographic and cultural context and with similar DCE valuation methodology. The cancer patient value set—with a broader range and lower

utility values—reflects a key finding from DCE tasks: cancer patients assigned higher priority to improvements in QLU-C10D attributes than to survival time extension alone. Moreover, the cancer patient value set exhibits greater sensitivity in distinguishing between different disease subgroups, perhaps better capturing the prioritization of symptom control and quality of life improvements. These insights provide valuable guidance for selecting reference groups in health economic evaluation and underscore the importance of tailoring value sets to patient populations in specific contexts.

Among Chinese cancer patients, physical functioning, pain, and emotional functioning exhibit the highest utility decrements, which is consistent with findings from Singaporean cancer patients [50]. This similarity is likely influenced by shared Southeast Asian cultural contexts. However, in the value set for Chinese general population, emotional functioning exhibits the smallest utility decrement [45]. This discrepancy may be attributed to China's unique disease culture and traditional Chinese medicine belief that "emotions can cause cancer, and emotions can heal cancer" [68]. Utility decrements from the cancer patient value set exceed those from the general population across most dimensions, particularly symptom-related ones. The consequence is that the cancer patient health state values are generally lower than those of the general population, with the maximum difference at the pit state (-0.092 versus 0.083). The most likely explanation of preference differences that patients' have a nuanced understanding of health impairments through their lived experience and conversely that the general population tend to underestimate the impact of symptoms due to a lack of firsthand experience [9, 69]. It is important to note that cross-population differences could also reflect scale heterogeneity (i.e., variations in choice consistency) [70], as well as differences in model constraints applied to address non-monotonicities between groups. These factors may influence utility estimates beyond true preference differences, warranting cautious interpretation. While some studies suggest that patients may evaluate the impact of health impairments less than the general population [12, 13], this is often attributed to generic PBMs failing to capture patients' specific concerns about symptoms and quality of life. A systematic review found that preference differences between patients and the general population are in the same direction as in our study, and are more pronounced in disease-specific PBMs than in generic PBMs [5].

Compared to the value set of the general population, cancer patients demonstrate significantly larger utility decrements for lower severity levels (Level 2) symptoms, such as fatigue, nausea, and bowel problems, which is similar with findings from the Singaporean cancer patient value set [50]. This leads to more pronounced differences between patient versus general population value sets for better health states, while differences are smaller for worse health states. It can be inferred that cancer patients prioritize alleviating mild symptoms, likely due to their heightened sensitivity to physical vulnerability and the cumulative burden of long-term treatments. Physiologically, persistent chronic pain and treatment-related side effects significantly amplify patients' sensitivity to minor symptoms in cancer patients [71]. Psychologically, the profound anxiety about health deterioration triggered by a cancer diagnosis further intensifies this heightened sensitivity[72]. Our finding that differences in decrements between patient versus general population value sets were less pronounced at higher symptom severity levels was possibly due to patients' stronger desire for life extension and greater tolerance for discomfort [73]. This is supported by the higher time coefficient (0.882) observed in this patient cohort compared to the general population's time coefficient (0.531) [45]. A similar trend was evident in Singapore, where the time coefficient was notably higher for cancer patients (0.529) than for the general population (0.285)[50].

This study demonstrated the sensitivity of both value sets to cancer stage and performance status. However, when extrapolating intergroup differences to QALY gains based on changes in health states,

the general population value set would tend to yield lower QALY gains than the cancer patient value set due to the systematically higher utility values from the general population. In the absence of specific decision-making preferences, and in cancer-specific decision-making, the cancer patient value set may be preferred as it more accurately reflects the treatment priorities of the patients. Some authors argue that value sets from the general population may be more suitable for HTA agencies adopting a payer perspective, while patient-derived index scores are better suited for patient registries, population health studies, and personalized medicine [36]. Furthermore, the differences between this study's cancer patient value set and Singapore's highlight the need for country-specific cancer value sets. Although recruiting cancer patients and conducting face-to-face interviews are more costly, hybrid methods, such as recalibrating general population value sets from the patient perspective, could enhance QALY comparability. Incorporating patient preferences into value sets and reprioritizing items based on their perspectives [17] could also be applied to QLU-C10D preference studies, enabling flexible integration of both societal and patient-based value sets in unified evaluations. Additionally, the online elicitation of personal utility functions (OPUF) tool offers a feasible approach for constructing value sets based on the preferences of small samples of cancer patients, which warrants further exploration and validation in future research[74].

This study has notable strengths, including adherence to the internationally standardized MAUCa Consortium protocol, amended for the patient context by use of a trained interviewer to facilitate the DCE valuation task, as done in the QLU-C10D pilot patient valuation [10]. Cognitive interviews conducted in the pilot patient valuation revealed patients' need for additional effort, time and explanation to make the DCE tasks easier to understand. Our interviewers provided that extra time and explanation, resulting in more positive participant feedback on the valuation process relative to the pilot [10]. Another study strength was that it enabled comparisons of health preferences between Chinese cancer patients and general population, as well as cross-country comparisons. However, several limitations should be acknowledged. Firstly, patient recruitment was limited to two tertiary public hospitals in Harbin, China. While this approach ensured data quality and feasibility, it may limit the generalizability of our findings, as the recruited sample might not fully represent the broader cancer patient population across China in terms of socioeconomic, regional, and clinical diversity. Secondly, voluntary participation may have introduced sample bias, potentially favoring individuals with milder conditions or better treatment outcomes, while close interviewer involvement could have induced some social desirability bias. Thirdly, we did not apply education weighting (as used in the general population value set) due to inherent differences in education distributions between cancer patients and the general population; this omission may weaken the stability of extrapolating our findings for cross-group comparisons. Finally, as with all stated preference methods, our use of discrete choice experiments may be subject to hypothetical bias, where participants' stated choices may not perfectly reflect real-world decision-making.

5 Conclusion

This study developed the first QLU-C10D value set based on the preferences of Chinese cancer patients, offering a vital resource for patient-centered clinical decision-making in the country with the highest cancer incidence worldwide. Furthermore, by comparing value sets derived from cancer patients and the general population of the same country, language and culture, and using a very similar valuation protocol, the study underscores the potential differences between the two, highlighting the importance of more comprehensively integrating patient perspectives into existing evaluation frameworks.

References

1. Neumann, P.J., Goldie, S.J., Weinstein, M.C.: Preference-based measures in economic evaluation in health care. Annu Rev Public Health (2000). https://doi.org/10.1146/annurev.publhealth.21.1.587
2. Bertram, M.Y., Lauer, J.A., Stenberg, K., Edejer, T.T.T.: Methods for the Economic Evaluation of Health Care Interventions for Priority Setting in the Health System: An Update From WHO CHOICE. Int J Health Policy Manag (2021). https://doi.org/10.34172/ijhpm.2020.244
3. Cost-Effectiveness in Health and Medicine: Cost-Effectiveness in Health and Medicine. Oxford University Press (2016). https://doi.org/10.1093/acprof:oso/9780190492939.001.0001
4. Dolders, M.G., Zeegers, M.P., Groot, W., et al.: A meta-analysis demonstrates no significant differences between patient and population preferences. J Clin Epidemiol (2006). https://doi.org/10.1016/j.jclinepi.2005.07.020
5. Greiner, W., Klose, K.: Valuation of health-related quality of life and utilities in health economics. Z Evid Fortbild Qual Gesundhwes (2014). https://doi.org/10.1016/j.zefq.2014.02.004
6. Rowen, D., Azzabi Zouraq, I., Chevrou-Severac, H., van Hout, B.: International Regulations and Recommendations for Utility Data for Health Technology Assessment. PharmacoEconomics (2017). https://doi.org/10.1007/s40273-017-0544-y
7. IQWiG: https://www.iqwig.de/en/.
8. Dental and Pharmaceutical Benefits Agency (Tandvårds- och läkemedelsförmånsverket, TLV): https://www.government.se/government-agencies/dental-and-pharmaceutical-benefits-agency-tandvards--och-lakemedelsformansverket-tlv/
9. Badia, X., Herdman, M., Kind, P.: The influence of ill-health experience on the valuation of health. Pharmacoeconomics (1998). https://doi.org/10.2165/00019053-199813060-00005
10. Gamper, E.M., King, M.T., Norman, R., Loth, F.L.C., Holzner, B., Kemmler, G.: The EORTC QLU-C10D discrete choice experiment for cancer patients: a first step towards patient utility weights. J Patient Rep Outcomes (2022). https://doi.org/10.1186/s41687-022-00430-5
11. Brazier, J., Akehurst, R., Brennan, A., Dolan, P., Claxton, K., McCabe, C., Sculpher, M., Tsuchyia, A.: Should patients have a greater role in valuing health states? Appl Health Econ Health Policy (2005). https://doi.org/10.2165/00148365-200504040-00002
12. Noel, C.W., Lee, D.J., Kong, Q., Xu, W., Simpson, C., Brown, D., Gilbert, R.W., Gullane, P.J., Irish, J.C., Huang, S.H., O'Sullivan, B., Goldstein, D.P., de Almeida, J.R.: Comparison of Health State Utility Measures in Patients With Head and Neck Cancer. JAMA Otolaryngol Head Neck Surg (2015). https://doi.org/10.1001/jamaoto.2015.1314
13. Gries, K.S., Regier, D.A., Ramsey, S.D., Patrick, D.L.: Preferences for Prostate Cancer Outcomes: A Comparison of the Patient Perspective, the General Population Perspective, and a Population at Risk for Prostate Cancer. Value Health (2016). https://doi.org/10.1016/j.jval.2015.11.012
14. Peeters, Y., Stiggelbout, A.M.: Health state valuations of patients and the general public analytically compared: a meta-analytical comparison of patient and population health state utilities. Value Health (2010). https://doi.org/10.1111/j.1524-4733.2009.00610.x
15. Chai, Q., Yang, Z., Liu, X., An, D., Du, J., Ma, X., Rand, K., Wu, B., Luo, N., Schulenburg, J.M.G.V.D.: Valuation of EQ-5D-5L health states from cancer patients' perspective: a feasibility study. European Journal of Health Economics (2024).
16. Mihir, G., Ru, S., Tan, R., Ng, S., Pin, & Choo: Comparison of health state values derived from patients and individuals from the general population. Quality of Life Research An International

Journal of Quality of Life Aspects of Treatment Care & Rehabilitation (2017).
17. Ludwig, K., Ramos-Goñi, J.M., Oppe, M., Kreimeier, S., Greiner, W.: To What Extent Do Patient Preferences Differ From General Population Preferences? Value Health (2021). https://doi.org/10.1016/j.jval.2021.02.012
18. King, M.T., Costa, D.S., Aaronson, N.K., Brazier, J.E., Cella, D.F., Fayers, P.M., Grimison, P., Janda, M., Kemmler, G., Norman, R., Pickard, A.S., Rowen, D., Velikova, G., Young, T.A., Viney, R.: QLU-C10D: a health state classification system for a multi-attribute utility measure based on the EORTC QLQ-C30. Qual Life Res (2016). https://doi.org/10.1007/s11136-015-1217-y
19. Aaronson, N.K., Ahmedzai, S., Bergman, B., Bullinger, M., Cull, A., Duez, N.J., Filiberti, A., Flechtner, H., Fleishman, S.B., de Haes, J.C., et al.: The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst (1993). https://doi.org/10.1093/jnci/85.5.365
20. Giesinger, J.M., Efficace, F., Aaronson, N., Calvert, M., Kyte, D., Cottone, F., Cella, D., Gamper, E.M.: Past and Current Practice of Patient-Reported Outcome Measurement in Randomized Cancer Clinical Trials: A Systematic Review. Value Health (2021). https://doi.org/10.1016/j.jval.2020.11.004
21. Luo, N., Liu, G., Li, M., Guan, H., Jin, X., Rand-Hendriksen, K.: Estimating an EQ-5D-5L Value Set for China. Value Health (2017). https://doi.org/10.1016/j.jval.2016.11.016
22. Liu, G.G., Wu, H., Li, M., Gao, C., Luo, N.: Chinese time trade-off values for EQ-5D health states. Value Health (2014). https://doi.org/10.1016/j.jval.2014.05.007
23. Zhuo, L., Xu, L., Ye, J., Sun, S., Zhang, Y., Burstrom, K., Chen, J.: Time Trade-Off Value Set for EQ-5D-3L Based on a Nationally Representative Chinese Population Survey. Value Health (2018). https://doi.org/10.1016/j.jval.2018.04.1370
24. Liu, G.G., Guan, H., Jin, X., Zhang, H., Vortherms, S.A., Wu, H.: Rural population's preferences matter: a value set for the EQ-5D-3L health states for China's rural population. Health Qual Life Outcomes (2022). https://doi.org/10.1186/s12955-022-01917-x
25. Yang, Z., Jiang, J., Wang, P., Jin, X., Wu, J., Fang, Y., Feng, D., Xi, X., Li, S., Jing, M., Zheng, B., Huang, W., Luo, N.: Estimating an EQ-5D-Y-3L Value Set for China. Pharmacoeconomics (2022). https://doi.org/10.1007/s40273-022-01216-9
26. Wu, J., Xie, S., He, X., Chen, G., Bai, G., Feng, D., Hu, M., Jiang, J., Wang, X., Wu, H., Wu, Q., Brazier, J.E.: Valuation of SF-6Dv2 Health States in China Using Time Trade-off and Discrete-Choice Experiment with a Duration Dimension. Pharmacoeconomics (2021). https://doi.org/10.1007/s40273-020-00997-1
27. Pan, C.W., He, J.Y., Zhu, Y.B., Zhao, C.H., Luo, N., Wang, P.: Comparison of EQ-5D-5L and EORTC QLU-C10D utilities in gastric cancer patients. Eur J Health Econ (2023). https://doi.org/10.1007/s10198-022-01523-0
28. Bulamu, N.B., Vissapragada, R., Chen, G., Ratcliffe, J., Mudge, L.A., Smithers, B.M., Isenring, E.A., Smith, L., Jamieson, G.G., Watson, D.I.: Responsiveness and convergent validity of QLU-C10D and EQ-5D-3L in assessing short-term quality of life following esophagectomy. Health Qual Life Outcomes (2021). https://doi.org/10.1186/s12955-021-01867-w
29. Klapproth, C.P., Fischer, F., Rose, M., Karsten, M.M.: Health state utility differed systematically in breast cancer patients between the EORTC QLU-C10D and the PROMIS Preference Score. J Clin Epidemiol (2022). https://doi.org/10.1016/j.jclinepi.2022.09.010
30. Shaw, J.W., Bennett, B., Trigg, A., DeRosa, M., Taylor, F., Kiff, C., Ntais, D., Noon, K., King, M.T.,

Cocks, K.: A Comparison of Generic and Condition-Specific Preference-Based Measures Using Data From Nivolumab Trials: EQ-5D-3L, Mapping to the EQ-5D-5L, and European Organisation for Research and Treatment of Cancer Quality of Life Utility Measure-Core 10 Dimensions. Value Health (2021). https://doi.org/10.1016/j.jval.2021.05.022

31. Cao, Y., Li, H., Cheng, L.J., King, M.T., Kemmler, G., Cella, D., Yu, H., Huang, W., Luo, N.: A comparison of measurement properties between EORTC QLU-C10D and FACT-8D in patients with hematological malignances. Health Econ Rev (2024). https://doi.org/10.1186/s13561-024-00560-0
32. Seyringer, S., Pilz, M.J., Jansen, F., Büttner, M., King, M.T., Norman, R., Kemmler, G., Nerich, V., Holzner, B., Bottomley, A., Gamper, E.M.: Cancer-specific utility instrument for health economic evaluations: A synopsis of the EORTC QLU-C10D user manual and current validity evidence. Eur J Cancer (2025). https://doi.org/10.1016/j.ejca.2025.115235
33. King, M.T., Viney, R., Simon Pickard, A., Rowen, D., Aaronson, N.K., Brazier, J.E., Cella, D., Costa, D.S.J., Fayers, P.M., Kemmler, G., McTaggart-Cowen, H., Mercieca-Bebber, R., Peacock, S., Street, D.J., Young, T.A., Norman, R., Consortium, M.A.: Australian Utility Weights for the EORTC QLU-C10D, a Multi-Attribute Utility Instrument Derived from the Cancer-Specific Quality of Life Questionnaire, EORTC QLQ-C30. Pharmacoeconomics (2018). https://doi.org/10.1007/s40273-017-0582-5
34. Kemmler, G., Gamper, E., Nerich, V., Norman, R., Viney, R., Holzner, B., King, M., European Organisation for, R., Treatment of Cancer Quality of Life, G.: German value sets for the EORTC QLU-C10D, a cancer-specific utility instrument based on the EORTC QLQ-C30. Qual Life Res (2019). https://doi.org/10.1007/s11136-019-02283-w
35. McTaggart-Cowan, H., King, M.T., Norman, R., Costa, D.S.J., Pickard, A.S., Regier, D.A., Viney, R., Peacock, S.J.: The EORTC QLU-C10D: The Canadian Valuation Study and Algorithm to Derive Cancer-Specific Utilities From the EORTC QLQ-C30. MDM Policy Pract (2019). https://doi.org/10.1177/2381468319842532
36. Finch, A.P., Gamper, E., Norman, R., Viney, R., Holzner, B., King, M., Kemmler, G., Group, E.Q.o.L.: Estimation of an EORTC QLU-C10 Value Set for Spain Using a Discrete Choice Experiment. Pharmacoeconomics (2021). https://doi.org/10.1007/s40273-021-01058-x
37. Jansen, F., Verdonck-de Leeuw, I.M., Gamper, E., Norman, R., Holzner, B., King, M., Kemmler, G., European Organisation for, R., Treatment of Cancer Quality of Life, G.: Dutch utility weights for the EORTC cancer-specific utility instrument: the Dutch EORTC QLU-C10D. Qual Life Res (2021). https://doi.org/10.1007/s11136-021-02767-8
38. Nerich, V., Gamper, E.M., Norman, R., King, M., Holzner, B., Viney, R., Kemmler, G.: French Value-Set of the QLU-C10D, a Cancer-Specific Utility Measure Derived from the QLQ-C30. Appl Health Econ Health Policy (2021). https://doi.org/10.1007/s40258-020-00598-1
39. Revicki, D.A., King, M.T., Viney, R., Pickard, A.S., Mercieca-Bebber, R., Shaw, J.W., Muller, F., Norman, R.: United States Utility Algorithm for the EORTC QLU-C10D, a Multiattribute Utility Instrument Based on a Cancer-Specific Quality-of-Life Instrument. Medical decision making : an international journal of the Society for Medical Decision Making (2021). https://doi.org/10.1177/0272989X211003569
40. Lehmann, J., Rojas-Concha, L., Petersen, M.A., Holzner, B., Norman, R., King, M.T., Kemmler, G., Group, E.Q.o.L.: Danish value sets for the EORTC QLU-C10D utility instrument. Qual Life Res (2024). https://doi.org/10.1007/s11136-023-03569-w
41. Shiroiwa, T., King, M.T., Norman, R., Muller, F., Campbell, R., Kemmler, G., Murata, T.,

Shimozuma, K., Fukuda, T.: Japanese value set for the EORTC QLU-C10D: A multi-attribute utility instrument based on the EORTC QLQ-C30 cancer-specific quality-of-life questionnaire. Qual Life Res (2024). https://doi.org/10.1007/s11136-024-03655-7

42. Rohde, G., Lehmann, J., Pilz, M.J., Rojas-Concha, L., Holzner, B., King, M.T., Norman, R., Kemmler, G.: Norwegian and Swedish value sets for the EORTC QLU-C10D utility instrument. Quality of Life Research. https://doi.org/10.1007/s11136-024-03824-8
43. Norman, R., Mercieca-Bebber, R., Rowen, D., Brazier, J.E., Cella, D., Pickard, A.S., Street, D.J., Viney, R., Revicki, D., King, M.T.: U.K. utility weights for the EORTC QLU-C10D. Health Econ (2019). https://doi.org/10.1002/hec.3950
44. Gamper, E.M., King, M.T., Norman, R., Efficace, F., Cottone, F., Holzner, B., Kemmler, G.: EORTC QLU-C10D value sets for Austria, Italy, and Poland. Qual Life Res (2020). https://doi.org/10.1007/s11136-020-02536-z
45. Cao, Y., Xu, J., Norman, R., King, M.T., Kemmler, G., Huang, W., Luo, N.: Chinese utility weights for the EORTC cancer-specific utility instrument QLU-C10D. Qual Life Res (2024). https://doi.org/10.1007/s11136-024-03776-z
46. Jansen, F., Coupé, V.M.H., Eerenstein, S.E.J., Cnossen, I.C., van Uden-Kraan, C.F., de Bree, R., Doornaert, P., Halmos, G.B., Hardillo, J.A.U., van Hinte, G., Honings, J., Leemans, C.R., Verdonck-de Leeuw, I.M.: Cost-utility and cost-effectiveness of a guided self-help head and neck exercise program for patients treated with total laryngectomy: Results of a multi-center randomized controlled trial. Oral Oncol (2021). https://doi.org/10.1016/j.oraloncology.2021.105306
47. Yang, M., Vioix, H., Hook, E.S., Hatswell, A.J., Batteson, R.L., Gaumond, B.R., O'Brate, A., Popat, S., Paik, P.K.: Health Utility Analysis of Tepotinib in Patients With Non-Small Cell Lung Cancer Harboring MET Exon 14 Skipping. Value Health (2023). https://doi.org/10.1016/j.jval.2023.02.007
48. Kerba, M., Lourenco, R.A., Sahgal, A., Cardet, R.F., Siva, S., Ding, K., Myrehaug, S.D., Masucci, G.L., Brundage, M., Parulekar, W.R.: An Economic Analysis of SC24 in Canada: A Randomized Study of SBRT Compared With Conventional Palliative RT for Spinal Metastases. Int J Radiat Oncol Biol Phys (2024). https://doi.org/10.1016/j.ijrobp.2023.12.052
49. Byrne, A., De Abreu Lourenco, R., Govindaraj, R., Ball, D., Le, H.: A cost-effectiveness analysis of stereotactic ablative radiotherapy versus conventionally fractionated radiotherapy in the management of stage 1 non-small-cell lung cancer: Results from the TROG 09.02 CHISEL study. J Med Imaging Radiat Oncol (2024). https://doi.org/10.1111/1754-9485.13755
50. Gandhi, M., Kanesvaran, R., Rashid, M.F.B.H., Chong, D.Q., Chay, W.Y., Tan, L.Y., Norman, R., King, M.T., Luo, N.: Valuation of the EORTC Quality of Life Utility Core 10 Dimensions (QLU-C10D) in a Multi-ethnic Asian Setting: How Does Having Cancer Matter? Pharmacoeconomics (2024). https://doi.org/10.1007/s40273-024-01432-5
51. Han, B., Zheng, R., Zeng, H., Wang, S., Sun, K., Chen, R., Li, L., Wei, W., He, J.: Cancer incidence and mortality in China, 2022. J Natl Cancer Cent (2024). https://doi.org/10.1016/j.jncc.2024.01.006
52. Brown, R.C., Heines, S., Witt, A., Braehler, E., Fegert, J.M., Harsch, D., Plener, P.L.: The impact of child maltreatment on non-suicidal self-injury: data from a representative sample of the general population. BMC Psychiatry (2018). https://doi.org/10.1186/s12888-018-1754-3
53. Kessler, R.C., Andrews, G., Colpe, L.J., Hiripi, E., Mroczek, D.K., Normand, S.L., Walters, E.E., Zaslavsky, A.M.: Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med (2002). https://doi.org/10.1017/s0033291702006074
54. Rabin, R., de Charro, F.: EQ-5D: a measure of health status from the EuroQol Group. Ann Med

(2001). https://doi.org/10.3109/07853890109002087
55. Mukuria, C., Rowen, D., Mulhern, B., McDool, E., Kharroubi, S., Bjorner, J.B., Brazier, J.E.: The Short Form 6 Dimensions (SF-6D): Development and Evolution. Appl Health Econ Health Policy. https://doi.org/10.1007/s40258-024-00919-8
56. Oken, M.M., Creech, R.H., Tormey, D.C., Horton, J., Davis, T.E., McFadden, E.T., Carbone, P.P.: Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol (1982).
57. Norman, R., Viney, R., Aaronson, N.K., Brazier, J.E., Cella, D., Costa, D.S., Fayers, P.M., Kemmler, G., Peacock, S., Pickard, A.S., Rowen, D., Street, D.J., Velikova, G., Young, T.A., King, M.T.: Using a discrete choice experiment to value the QLU-C10D: feasibility and sensitivity to presentation format. Qual Life Res (2016). https://doi.org/10.1007/s11136-015-1115-3
58. Norman, R., Kemmler, G., Viney, R., Pickard, A.S., Gamper, E., Holzner, B., Nerich, V., King, M.: Order of Presentation of Dimensions Does Not Systematically Bias Utility Weights from a Discrete Choice Experiment. Value Health (2016). https://doi.org/10.1016/j.jval.2016.07.003
59. Gamper, E.M., Holzner, B., King, M.T., Norman, R., Viney, R., Nerich, V., Kemmler, G.: Test-Retest Reliability of Discrete Choice Experiment for Valuations of QLU-C10D Health States. Value Health (2018). https://doi.org/10.1016/j.jval.2017.11.012
60. Roudijk, B., Donders, A.R.T., Stalmeier, P.F.M.: Setting Dead at Zero: Applying Scale Properties to the QALY Model. Med Decis Making (2018). https://doi.org/10.1177/0272989x18765184
61. Miyamoto, J.M., Wakker, P.P., Bleichrodt, H., Peters, H.J.M.: The Zero-Condition: A Simplifying Assumption in QALY Measurement and Multiattribute Utility. Management Science (1998).
62. Norman, R., Cronin, P., Viney, R.: A pilot discrete choice experiment to explore preferences for EQ-5D-5L health states. Appl Health Econ Health Policy (2013). https://doi.org/10.1007/s40258-013-0035-z
63. Cicchetti, Domenic, V.: Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology. Psychological Assessment (1994).
64. Cohen, J.: A power primer. Psychol Bull (1992). https://doi.org/10.1037//0033-2909.112.1.155
65. Zeng, X., Sui, M., Liu, B., Yang, H., Liu, R., Tan, R.L., Xu, J., Zheng, E., Yang, J., Liu, C., Huang, W., Yu, H., Luo, N.: Measurement Properties of the EQ-5D-5L and EQ-5D-3L in Six Commonly Diagnosed Cancers. Patient (2021). https://doi.org/10.1007/s40271-020-00466-z
66. Yu, H., Zeng, X., Sui, M., Liu, R., Tan, R.L., Yang, J., Huang, W., Luo, N.: A head-to-head comparison of measurement properties of the EQ-5D-3L and EQ-5D-5L in acute myeloid leukemia patients. Qual Life Res (2021). https://doi.org/10.1007/s11136-020-02644-w
67. Pickard, D.A.S., Wilke, C.T., Lin, H.W., Lloyd, A.: Health Utilities Using the EQ-5D in Studies of Cancer. Pharmacoeconomics (2007).
68. Scheid, V.: Depression, constraint, and the liver: (Dis)assembling the treatment of emotion-related disorders in Chinese medicine. Cult Med Psychiatry (2013). https://doi.org/10.1007/s11013-012-9290-y
69. Gandhi, M., Tan, R.S., Ng, R., Choo, S.P., Chia, W.K., Toh, C.K., Lam, C., Lee, P.T., Latt, N.K.Z., Rand-Hendriksen, K., Cheung, Y.B., Luo, N.: Comparison of health state values derived from patients and individuals from the general population. Qual Life Res (2017). https://doi.org/10.1007/s11136-017-1683-5
70. Vass, C.M., Wright, S., Burton, M., Payne, K.: Scale Heterogeneity in Healthcare Discrete Choice Experiments: A Primer. Patient (2018). https://doi.org/10.1007/s40271-017-0282-4
71. Chen, X., Wang, Z., Zhou, J., Li, Q.: Intervention and coping strategies for self-perceived burden of

patients with cancer: A systematic review. Asia Pac J Oncol Nurs (2023). https://doi.org/10.1016/j.apjon.2023.100231
72. Versluis, M.A.J., van de Poll-Franse, L.V., Zijlstra, M., van Laarhoven, H.W.M., Vreugdenhil, G., Henselmans, I., Brom, L., Kuip, E.J.M., van der Linden, Y.M., Raijmakers, N.H.J.: Changes in perception of prognosis in the last year of life of patients with advanced cancer and its associated factors: Longitudinal results of the eQuiPe study. Palliat Med (2024). https://doi.org/10.1177/02692163241301220
73. Audrey, S., Abel, J., Blazeby, J.M., Falk, S., Campbell, R.: What oncologists tell patients about survival benefits of palliative chemotherapy and implications for informed consent: qualitative study. Bmj (2008). https://doi.org/10.1136/bmj.a752
74. Schneider, P., Hout, B.V., Heisen, M., Brazier, J., Devlin, N.: The Online Elicitation of Personal Utility Functions (OPUF) tool: a new method for valuing health states. Wellcome Open Research (2022).

Supplementary material

Quality of Life Utility-Core 10 Dimensions (QLU-C10D) scores based on cancer patients' health preferences are more sensitive than those based on general public's health preferences

Yiyin Cao[1]*, Haofei Li[1]*, Xu Jin[1], Madeleine T. King[2], Richard Norman[3], Ling Jie Cheng[4,5,6], Hongjuan Yu[7#], Weidong Huang[1#], Nan Luo[6]

[1] School of Health Management, Harbin Medical University, Harbin, China

[2] School of Psychology, University of Sydney, Sydney, Australia

[3] School of Public Health, Curtin University, Perth, Australia

[4] Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom

[5] Yong Loo Lin School of Medicine, National University of Singapore, Singapore

[6] Saw Swee Hock School of Public Health, National University of Singapore, Singapore

[7] Department of Hematology, The First Affiliated Hospital of Harbin Medical University, Harbin, China

*These authors have made equal contributions and shared first authorship.

[#]Correspondence:

Weidong Huang: huangweidong@hrbmu.edu.cn

Hongjuan Yu:yuhongjuan2008@163.com

Contents

Table A Health state classification system of the QLU-C10D 20

Fig. A Example of a choice set used in the discrete choice experiment valuation task to determine utility weights for the QLU-C10D 22

Fig. B An example choice set from the discrete choice experiment valuation task in Chinese as presented to survey participants 23

Fig. C Results from the QLU-C10D valuation survey feedback questions from patients 24

Data Quality Results and Discussion 25

Fig. D Distribution of time to completion by choice set **25**

Fig. E Relationship of completion time decile with model fit **26**

Fig. F Utility values of the Chinese cancer patients who participated in this study, calculated based on value sets derived from those cancer patients and the Chinese general population 26

Fig. G Bland-Altman plot of utility values of the Chinese cancer patients who participated in this study calculated based on value sets derived from those cancer patients and the Chinese general population 27

Table A Health state classification system of the QLU-C10D

Dimension	Level	Stem	Descriptor	QLQ-C30 item scores
Physical functioning	1	You have…	No trouble taking a long walk outside of the house	Item 2 (long walk) = 1
	2		No trouble taking a short walk outside of the house, but at least a little trouble taking a long walk	Item 3 (short walk) = 1 AND Item 2 ≥ 2
	3		At least a little trouble taking a short walk outside of the house, and at least a little trouble taking a long walk	Item 3 = 2 AND Item 2 ≥ 2
	4		Quite a bit or very much trouble taking a short walk outside the house	Item 3 ≥ 3 AND Item 2 ≥ 2
Role functioning	1	You are limited in pursuing your work or other daily activities…	Not at all	Item 6 = 1
	2		A little	Item 6 = 2
	3		Quite a bit	Item 6 = 3
	4		Very much	Item 6 = 4
Social functioning	1	Your physical condition or medical treatment interferes with your social or family life…	Not at all	Items 26 AND 27 = 1
	2		A little	Items 26 OR 27 = 2
	3		Quite a bit	Items 26 OR 27 = 3
	4		Very much	Items 26 OR 27 = 4
Emotional functioning	1	You feel depressed…	Not at all	Item 24 = 1
	2		A little	Item 24 = 2
	3		Quite a bit	Item 24 = 3
	4		Very much	Item 24 = 4
Pain	1	You have pain…	Not at all	Item 9 = 1
	2		A little	Item 9 = 2
	3		Quite a bit	Item 9 = 3
	4		Very much	Item 9 = 4
Fatigue	1	You feel tired…	Not at all	Item 18 = 1
	2		A little	Item 18 = 2
	3		Quite a bit	Item 18 = 3

	4		Very much	Item 18 = 4
Sleep	1	You have trouble sleeping…	Not at all	Item 11 = 1
	2		A little	Item 11 = 2
	3		Quite a bit	Item 11 = 3
	4		Very much	Item 11 = 4
Appetite	1	You lack appetite…	Not at all	Item 13 = 1
	2		A little	Item 13 = 2
	3		Quite a bit	Item 13 = 3
	4		Very much	Item 13 = 4
Nausea	1	You feel nauseated…	Not at all	Item 14 = 1
	2		A little	Item 14 = 2
	3		Quite a bit	Item 14 = 3
	4		Very much	Item 14 = 4
Bowel problems	1	You…	do not have constipation or diarrhoea at all	Items 16 AND 17 = 1
	2		have a little constipation or diarrhoea	Items 16 OR 17 = 2
	3		have constipation or diarrhoea quite a bit	Items 16 OR 17 = 3
	4		have constipation or diarrhoea very much	Items 16 OR 17 = 4

	Situation A	Situation B
In taking a long walk	You have a little trouble	You have a little trouble
In taking a short walk	You have a little trouble	You have a little trouble
You are limiting in pursuing your daily activities	Quite a bit	Quite a bit
Your physical condition interferes with your social or family life	Quite a bit	A little
You feel depressed	Quite a bit	Quite a bit
You have pain	A little	Very much
You feel tired	A little	A little
You have trouble sleeping	Not at all	Not at all
You lack appetite	Quite a bit	Quite a bit
You feel nauseated	A little	Quite a bit
You have constipation or diarrhoea	Not at all	Quite a bit
You will live in this health state for	2 years and then die	5 years and then die
Which situation would you prefer?	○	○

Fig. A Example of a choice set used in the discrete choice experiment valuation task to determine utility weights for the QLU-C10D

	健康状态 A	健康状态 B
对于长距离步行	有一点困难	有一点困难
对于短距离步行	有一点困难	有一点困难
您的工作和生活受到疾病限制	相当多	相当多
您的身体状况或医学治疗干扰到您的家庭或社交活动	相当多	有一点
您感到沮丧	相当多	相当多
您感到疼痛	有一点	非常多
您感到疲劳	有一点	有一点
您睡眠困难	没有	没有
您胃口不好	相当多	相当多
您感到恶心	有一点	相当多
您有便秘或腹泻	没有	相当多
您将在这种健康状态下生活	2 年,然后去世	1 年,然后去世
您选择哪一个？	○	○

Fig. B An example choice set from the discrete choice experiment valuation task in Chinese as presented to survey participants

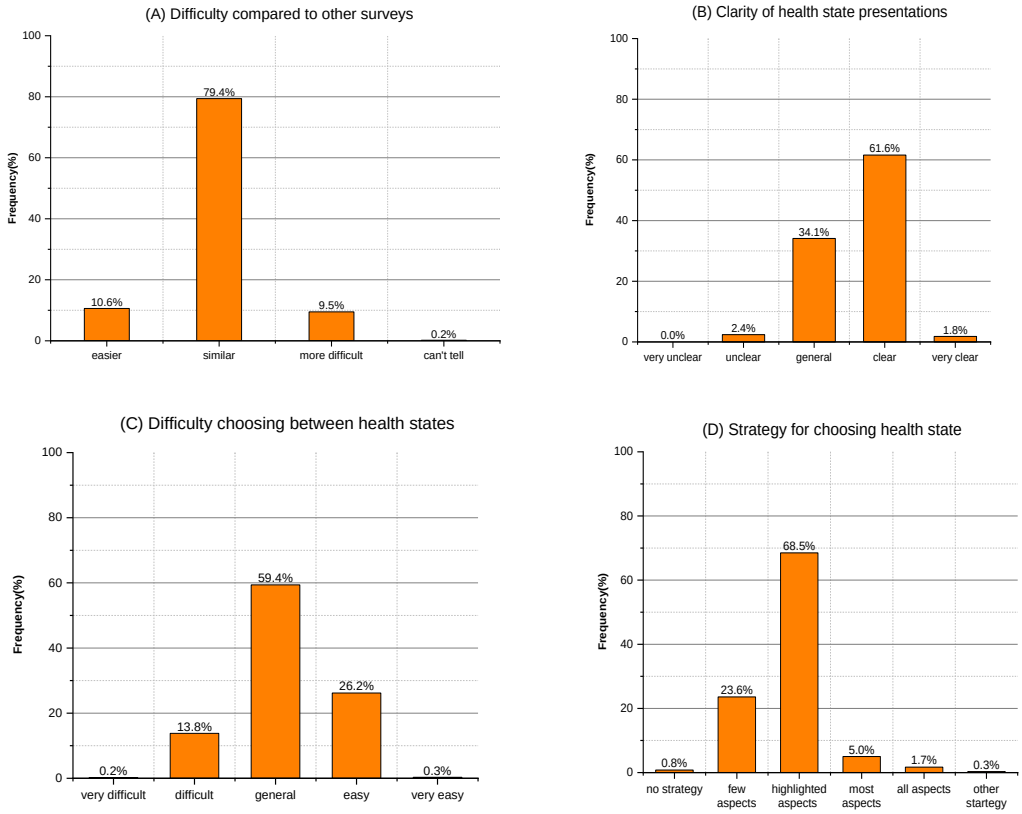

Fig. C Results from the QLU-C10D valuation survey feedback questions from patients

Data Quality Results and Discussion

Overall, as respondents became more familiar with the choice task, all respondents across the completion time deciles sped up (Figure D). Respondents in the sixth decile yielded the most statistically significant coefficients (24 of a possible 31). Although faster respondents produced fewer statistically significant coefficients, suggesting more random data, the pseudo-R^2 values of the fastest decile groups were relatively high. The slower respondents also produced fewer statistically significant coefficients, and the two slowest two deciles had the lowest pseudo-R^2 values (Figure E).

This phenomenon may be related to the following differential patterns of survey engagement in this cancer patient cohort. During the interviews, we observed that younger and healthier patients were often more willing to complete the questionnaire independently, with less need for the interviewer's explanation and interaction. They tended to complete tasks more quickly, making them more likely to appear in the earlier percentiles. In contrast, we observed that patients in poorer health or older patients generally required more time and assistance from the interviewer, even though both the interviewer's feedback and statistical results are less ideal. These informal observations were confirmed by corresponding differential composition of age and performance status across different percentiles (data not shown).

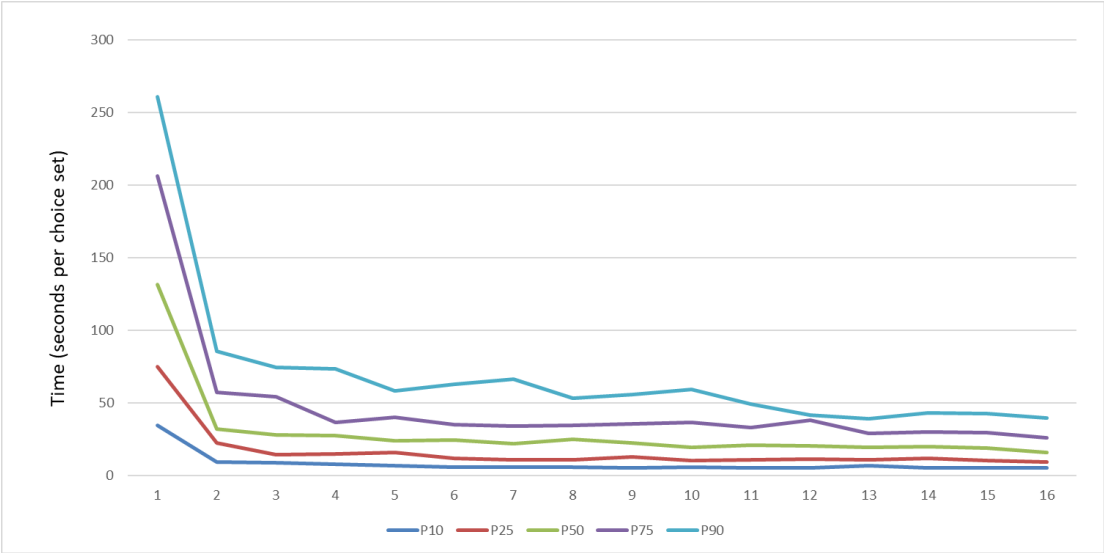

Fig. D Distribution of time to completion by choice set

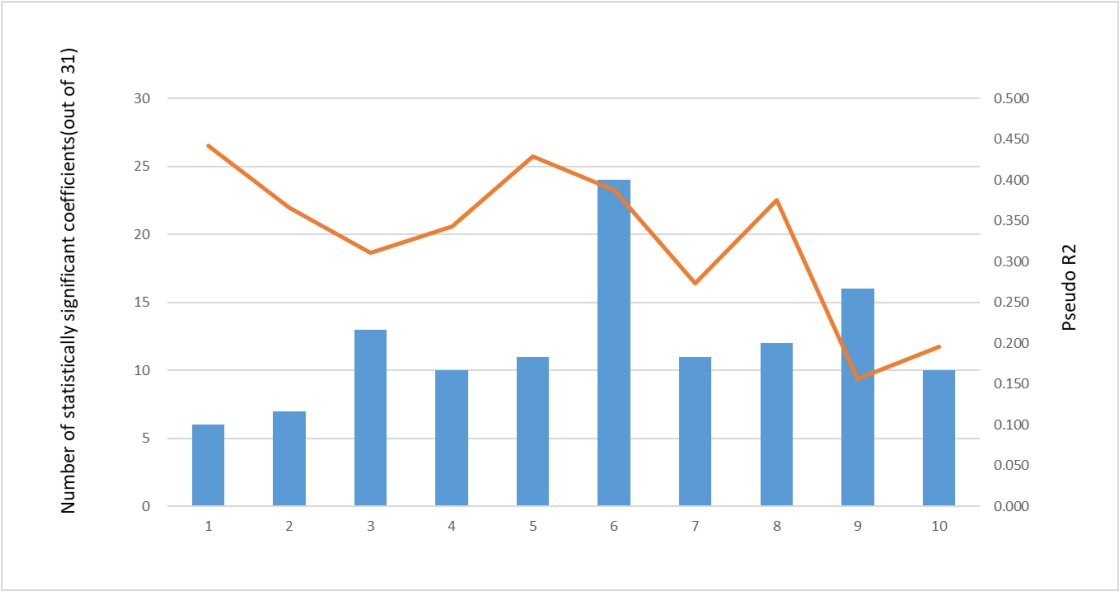

Fig. E **Relationship of completion time decile with model fit** (pseudo R-squared, orange line) and number of statistically significant coefficients (blue bars). Decile 1 contains the fastest 10% of respondents and decile 10 contains the slowest 10%.

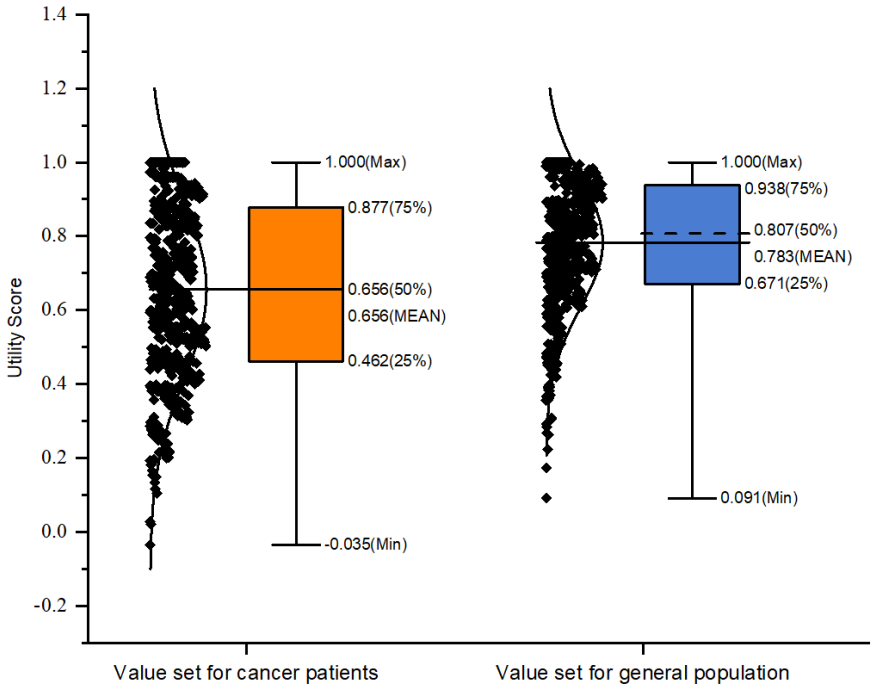

Fig. F Utility values of the Chinese cancer patients who participated in this study, calculated based on value sets derived from those cancer patients and the Chinese general population

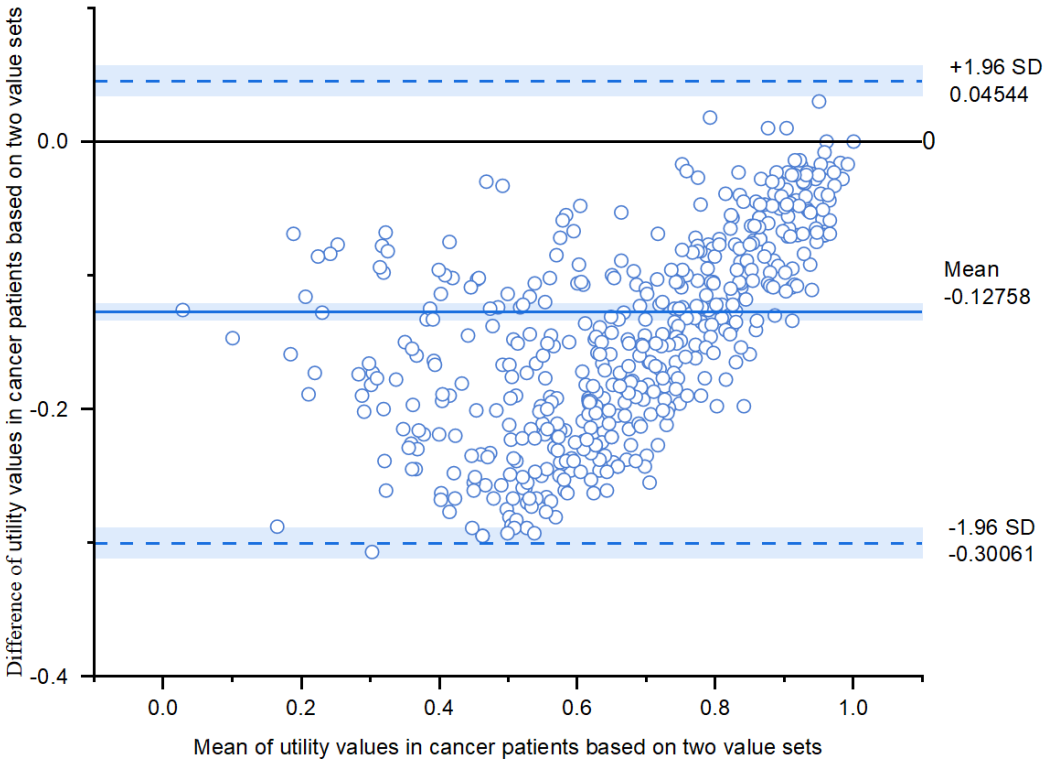

Fig. G Bland-Altman plot of utility values of the Chinese cancer patients who participated in this study calculated based on value sets derived from those cancer patients and the Chinese general population

Quality of Life Utility-Core 10 Dimensions (QLU-C10D) scores based on cancer patients' health preferences are more sensitive than those based on general public's health preferences

Abstract

Objective: This study aimed to develop a value set for the European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Utility-Core 10 Dimensions (QLU-C10D) based on the preferences of Chinese cancer patients, and to compare results with the Chinese general population value set.

Methods: Patients with mixed cancer diagnoses were recruited from two hospitals in China between September and November 2024. Patients completed the EORTC Core Quality of Life Questionnaire (QLQ-C30) followed by a discrete choice experiment (DCE) valuation survey consisting of 16 choice sets through face-to-face interviews. A conditional logit model was applied to analyze patient preferences. The characteristics and sensitivity of the derived QLU-C10D value set were compared with those of the Chinese general population QLU-C10D value set. Sensitivity was evaluated by ability to discriminate cancer stage and performance status, and effect sizes (ES) were estimated.

Results: Data from 659 cancer patients were analyzed. The cancer patients' value set showed a broader utility range (-0.092 to 1.0) compared to the general population's value set (0.083 to 1). The mean utility score for the cancer patient sample (0.656) was significantly lower than that for the general population (0.738, $P < 0.01$). Both value sets demonstrated good sensitivity, but the cancer patients' value set was more effective in distinguishing clinical subgroup differences (ES = 2.76–4.46 vs. ES = 2.47–3.48).

Conclusion: Differences in value sets between cancer patients and the general population underscore the importance of more comprehensively integrating patient perspectives into existing evaluation frameworks.

Keywords: Value Set, Cancer Patients, General Population, QLU-C10D, China

1 Introduction

Cost-utility analysis (CUA) is a key method for reimbursement decisions in healthcare, typically using preference-based health-related quality of life (HRQoL) utility values to derive quality-adjusted life years (QALYs). The QALY combines morbidity and mortality into a single metric, weighed by preferences that reflect the relative desirability of health states, serving as the basis for evaluating the impact of health interventions [1]. Country-specific value sets, reflecting societal preferences and collective benefits, are considered ideal for healthcare decision-making [2, 3]. This practice, rooted in the principle of maximizing societal value, has been widely endorsed by economic evaluation guidelines across multiple countries [4].

However, in the context of patient-centered decision-making, the debate over prioritizing general population versus patient-derived preferences remains unresolved [4-6]. While general population preferences capture societal benefits, patient preferences represent the lived experiences associated with specific health states. In Germany, the HTA framework led by the Institute for Quality and Efficiency in Health Care focuses on qualitative assessments of "added clinical benefit" from the patient perspective, such as survival gains and symptom improvement. It directly incorporates patient-reported outcomes (PROs) to evaluate therapies against comparators, ensuring patient voices are central to treatment benefit evaluations[7]. Similarly, Sweden's Dental and Pharmaceutical Benefits Agency emphasizes ethical principles like "human value and need" and "solidarity" in its health economic evaluations, which prioritize patient experiences. These evaluations integrate condition-specific PROs to align resource allocation with patients' actual needs [8]. By incorporating patient preferences, HTAs can better align resource allocation with actual patient needs, improve perceived health benefits and satisfaction, and enhance personalized healthcare [9-11]. Therefore, integrating value sets representing preferences informed by the lived experience alongside societal perspectives offers a more comprehensive framework for clinical and policy decision-making.

Significant preference differences highlight the critical importance of choosing between patient and general population preferences. A meta-analysis conducted in 2006 indicated that minimal or no significant differences between utility values derived from these groups [5]. However, recent studies increasingly reveal notable discrepancies between these groups, encouraging the development of patient-specific value sets. For example, some studies suggest that utility values derived from patients tend to be systematically higher [12, 13], while others report the opposite trend [14-16]. Moreover, the direction and magnitude of these differences appear to be influenced by factors such as disease type and national context [17]. These findings reflect the unique experiences, recovery expectations, and treatment perceptions of patients, highlighting the limitations of universally applying general population-based value sets in healthcare decision-making. This calls for further exploration into the role of patient preferences in health evaluations and their divergence from general preferences.

In recent years, the European Organization for Research and Treatment of Cancer (EORTC), in collaboration with the Multi-Attribute Utility in Cancer (MAUCa) Consortium, developed the QLU-C10D [18], a cancer-specific preference-based measure (PBM) derived from the widely used Quality of Life Questionnaire - Core 30 (QLQ-C30) [19, 20]. In China, value sets based on general population preferences have been developed for several generic PBMs, including the EQ-5D [21-24], EQ-5D-Y [25], and SF-6D [26]. While these PBMs are widely applied in HTA, they are not cancer-specific and may not fully capture the symptom burden and HRQoL changes experienced by cancer patients. The QLU-C10D, by contrast, has shown high sensitivity in detecting subtle health changes and distinguishing between

health states among cancer patients [27-32]. Value sets for the QLU-C10D, developed from general population preferences, are now available in 16 countries [33-45], including China, facilitating its application in HTA [46-49]. However, patient-derived QLU-C10D value sets have received increasing attention only recently. Gamper et al. were the first to apply the QLU-C10D valuation survey among cancer patients in Austria, confirming its feasibility and laying the groundwork for patient-specific value sets [10]. Gandhi et al. extended this work by comparing utilities derived from general population and cancer patient preferences in Singapore [50]. However, the impact of such differences on CUAs remains underexplored, warranting further research to advance personalized healthcare and refine economic evaluation methods.

This study aimed to develop QLU-C10D utility weights for cancer patients in China, the country with the highest global cancer incidence [51], to inform clinical decision-making. Additionally, it aimed to explore differences in Chinese QLU-C10D value sets based on health preferences from cancer patients versus the general population, and consider how these differences may impact CUA.

2 Methods

The research design and analytical approach in this study follow the established QLU-C10D valuation methods used in previous studies, which are based on general population preferences from multiple countries [33-45]. These methods generally adhere to the standardized EORTC protocol for value set endorsement, with two adjustments for the patient context: recruitment and administration of the valuation study. Previous studies have demonstrated the feasibility of conducting QLU-C10D valuations among cancer patients [10].

2.1 Sample recruitment and management

From September to November 2024, the study recruited hospitalized cancer patients from two Grade A tertiary public hospitals in Harbin, the capital of Heilongjiang Province, including a cancer specialty hospital. The aim was to ensure a diverse sample in terms of age, residence, diagnosis, and treatment modalities. Participants were selected by hospital doctors or nurses based on the following criteria: (1) confirmed clinical diagnosis of cancer; (2) anticipated survival of at least one year for patient safety and ethical compliance; (3) age ≥18 years; and (4) proficiency in reading and communicating in Chinese.

Eligible patients provided written informed consent and participated in face-to-face interviews conducted by trained interviewers in hospital wards. First, as a familiarization exercise, the patient self-completed the QLQ-C30, albeit with the items read to them by an interviewer. The interviewer then explained the discrete choice experiment (DCE) task, using example questions to ensure comprehension. Responses were recorded electronically, with interviewers providing clarifications when needed. Standardized training was implemented, and a protocol was strictly followed to minimize interviewer bias and ensure data quality.

2.2 The QLU-C10D health state classification

The scenarios in the DCE focused on the QLU-C10D dimensions, derived from the QLQ-C30 by the MAUCa Consortium (2010–2016) using item response theory, qualitative interviews with cancer patients, expert opinion of oncologists and HRQOL researchers [18]. The QLU-C10D includes 13 items across ten health dimensions—physical, role, social, and emotional functioning; pain; fatigue; sleep disorders; appetite; nausea; and bowel problems. Each dimension has four severity levels (reflecting the source QLQ-C30 items levels), defining 1,048,576 unique health states [18]. Table A (supplementary material) details the classification system.

2.3 Valuation survey and DCE design

The valuation survey followed the original methods designed by the MAUCa Consortium and EORTC QOL Group [33], modified only in recruitment and implementation methods to enable cancer patient involvement. The DCE design and data management were supported by Survey Engine, and complied with the International Code on Market, Opinion and Social Research and Data Analytics [52].

Participants completed a survey comprising four sections: (1) self-reported health scales (with assistance of interviewers for reading and entering responses only), including QLQ-C30 [19], Kessler Psychological Distress Scale (Kessler-10) [53] for mental health, EQ-5D-5L [54], and SF-6Dv2 [55], all using officially endorsed Chinese versions; (2) DCE tasks with 16 choice sets; (3) feedback on DCE tasks, covering task difficulty, clarity, and choice strategies; and (4) sociodemographic questions, such as gender, age, and residence. Clinical data, including diagnosis, cancer staging, treatment modalities, and ECOG performance status, were obtained from medical records [56].

This study employed the DCE-based QLU-C10Dvaluation protocol described by King et al. [33], widely used thereafter in valuation studies across numerous countries [33-45]. Patients completed 16 choice sets, selected from 960 combinations via a D-efficiency procedure (a key method under optimal design theory) that prioritizes maximizing the efficiency of main effect parameter estimates. Interaction effects were not included to balance survey feasibility and statistical precision, consistent with common practices [33-45]. To control biases, health state A/B assignment and the order of 11 attributes across 10 QLU-C10D domains (the physical functioning dimension was split into 2 attributes, one for 'long walk' and one for 'short walk' (see Table A from supplementary material) were randomized per respondent; a balanced incomplete block design (BIBD) defined four differing dimensions per set (others held equal) to minimize burden. Each set presented two hypothetical QLU-C10D health states, differing in four of ten HRQoL dimensions and survival times (1, 2, 5, or 10 years), with differences highlighted for clarity. Examples are shown in Figures A and B (Chinese version) in the supplementary material. Details on QLU-C10D methodology, including formats, attribute ordering, and reliability, are provided in related studies [10, 57-59].

2.4 Data quality control

First, face-to-face interviews were conducted to optimize survey comprehension and interactivity: trained interviewers clarified DCE tasks, offered assistance as required, and halted surveys exclusively under specific circumstances—including inability to understand the survey interface, persistent failure to grasp task requirements (despite repeated explanations), or voluntary requests to withdraw (e.g., due to fatigue). These situations were deemed to directly compromise data validity. A total of 7 surveys (1.06% of the initial sample) were terminated, resulting in a final sample of 659 participants. The primary reasons for termination were fatigue (n=5) and persistent comprehension difficulties (n=2). Notably, none of these terminations were due to uncertainty or difficulty in making choices during the choice tasks. Secondly, patients who provided the most negative responses to all four DCE feedback questions ("Survey is more difficult," "Health states very unclear," "Choices very difficult," and "No choice strategy") were excluded as part of a robustness check to ensure task comprehension. Lastly, to assess the impact of survey duration on data quality, respondents were divided into deciles by completion time. Pseudo-R^2 values and significant coefficients from conditional logit analysis were plotted to examine trends.

2.5 Data analysis

Descriptive statistics summarized the characteristics of the cancer patient sample; specifically,

frequency and percentage for categorical variables, mean and standard deviation for continuous variables.

Conditional logistic regression was employed to model utility and calculate utility weights by interacting QLU-C10D dimensions with survival time. Utility values were anchored at 0 ("dead") and 1 ("full health") to maintain ratio-scale properties[60, 61],yet remained unconstrained during initial estimation to capture states "worse than deaths". The utility of each option was calculated as follows:

$$U_{isj} = \alpha \mathrm{TIME}_{isj} + \beta X'_{isj} \mathrm{TIME}_{isj} + \varepsilon_{isj} \quad (1)$$

where TIME_{isj} represents the survival time presented in DCE option j, and X'_{isj} was a set of dummy variables corresponding to the levels of the health state in DCE option j. The error term, εisj, was assumed to follow a Gumbel distribution. Conditional logistic regression estimated parameters α (utility per life year) and β (utility weights for each health state level). Utility decrements were calculated as the ratio of β to α, reflecting the HRQoL-life expectancy trade-off. Non-monotonic coefficients were adjusted by constraining adjacent levels, as per standard practice[62].

Line charts were used to compare utility decrements across QLU-C10D dimensions and levels between the cancer patient value set versus the Chinese general population value set. The general population value set served as the baseline to illustrate the percentage changes in utility decrements for the cancer patient value set, visualized using bar charts. Additionally, 1,000 hypothetical health states were generated by randomly combining the 4 levels across the 10 dimensions of the QLU-C10D descriptive system. This sampling approach ensured proportional representation of each dimension's levels, with states spanning the full severity spectrum—from mild (e.g., most dimensions at levels 1–2) to severe (e.g., majority at levels 3–4)—and explicitly including the best-possible (all dimensions at level 1) and worst-possible (all dimensions at level 4) health states. This comprehensive coverage of potential health states was designed to ensure robust utility comparisons between the two value sets.

Box plots displayed utility scores derived from both value sets, highlighting ranges and means. Agreement was assessed using intraclass correlation coefficients (ICC; high > 0.75) and Bland-Altman plots [63]. Sensitivity analysis stratified patients by cancer stage (I–IV) and ECOG performance status (Grades 0–4). Mean utility scores for each subgroup were compared using one-way ANOVA. "Extreme subgroups" here refer to the two ends of the stratified variables: for cancer stage, these were Stage I (earliest) and Stage IV (most advanced); for ECOG performance status, they were Grade 0 (fully active) and Grade 4 (completely disabled). Effect sizes (ES) quantified the clinical significance of differences between these extreme subgroups, calculated as the mean score difference divided by their pooled standard deviation [64]. Incremental utility was defined as the absolute difference in mean utility values between these extreme subgroups, reflecting the utility change associated with moving from the most severe to the least severe end of the stratified variable. We hypothesized higher utility values for earlier cancer stages or lower ECOG grades[29, 65-67].

3 Results

3.1 Sample characteristics

Of the 727 participants, 68 were excluded due to missing medical records or incomplete disease information, leaving 659 patients with various cancer types. The most common cancers were lung, breast stomach, and thyroid. Cancers with a prevalence of less than 3% were grouped into the 'other' category, which accounted for 7% of the sample. The sample included 347 males (53%) and 312 females (47%), with 48% aged 60 or older, 61% residing in urban areas, 69% having a junior high education or below, and 91% married. Detailed socio-demographic characteristics are presented in Table 1.

Table 1 Socio-demographic characteristics of the study samples

	Cancer patient sample (n=659)	
	N	%
Gender[a]		
Male	347	52.7
Female	312	47.3
Age group		
18-39	53	8.0
40-59	287	43.6
≥60	319	48.4
Residence		
Urban	402	61.0
Rural	257	39.0
Education level		
Junior high school or below	454	68.9
Senior high school/technical school	118	17.9
Post-secondary	87	13.2
Marital status		
Single	21	3.2
Married	602	91.4
Widowed/divorced/separated	36	5.5
Cancer type		
Lung cancer	145	22.0
Breast cancer	108	16.4
Stomach cancer	87	13.2
Thyroid cancer	67	10.2
Lymphoma	53	8.0
Leukemia	46	7.0
Liver cancer	44	6.7
Colorectal cancer	33	5.0
Prostate cancer	28	4.2
Others	48	7.3

[a] No significant difference was found between the distribution in this study and the latest incidence rates reported by the National Cancer Center and the International Agency for Research on Cancer (IARC). (/www.iarc.who.int/wp-content/uploads/2024/02/pr345_E.pdf) ($\chi^2 = 0.343$, $p = 0.558$)

3.2 Data quality control

Most patients found the DCE survey difficulty comparable to other surveys (79%), with few citing unclear health state descriptions (2%) or difficulty choosing between health states (14%, "very difficult" 0.2%). A majority (69%) focused on the "highlighted aspects" of the tasks. Detailed feedback results are shown in Figure C (supplementary material). No participants were excluded for consistently negative responses.

The mean survey completion time was 29.95 minutes (range: 5.39–80.12). Completion times decreased with task familiarity (Figure D in the supplementary material). The median decile produced the highest number significant coefficients (25/31). The fastest decile produced the fewest significant coefficients (6/31) and yet achieved comparable or higher model fit (pseudo-R^2 value 0.442) than other deciles (Figure E in the supplementary material); for further detail and discussion see the supplementary material. These results confirm the dataset's high quality and reliability.

3.3 Utility estimates

As shown in the "Unconstrained" column of Table 2, incremental moves to worse levels within each dimension generally led to significantly greater coefficients. However, non-monotonicities were observed in role functioning levels 3 and 4, where level 4 was preferred over level 3. A likelihood-ratio

test confirmed no significant difference between these levels, so monotonicity was imposed and the model re-estimated. Utility estimates for the cancer patient sample after correction show similar utility weights in both unconstrained and monotonicity-corrected model.

Monotonicity-corrected results were used to develop the QLU-C10D value set for cancer patients in China. Utility scores are derived by mapping QLQ-C30 responses to the corresponding QLU-C10D dimensions and levels (Table A in the supplementary material). For example, for QLQ-C30 responses matching the QLU-C10D health state "3233412432," the utility score is calculated as: 1 - 0.275 - 0.052 - 0.035 - 0.057 - 0.125 - 0 - 0.046 - 0.083 - 0.075 - 0.030 = 0.222. The worst possible health state (PITS state 4444444444) had a utility value of −0.092, considerably lower than the Chinese general population PITS value (0.083)[45].

Table 2 QLU-C10D model coefficients and utility decrements for the cancer patient sample

		Unconstrained			Corrected for monotonicity		
Parameter	**Level[a]**	**Coefficient (α, β)**	**SE**	**Utility decrement (β/ α)**	**Coefficient (α, β)**	**SE**	**Utility decrement (β/ α)**
Time coefficient (α)	(linear)	0.885**	0.059		0.882 **	0.059	
Attributes (β)							
Physical Functioning	2	-0.145 **	0.026	-0.164	-0.145 **	0.027	-0.165
Physical Functioning	3	-0.242 **	0.032	-0.273	-0.242 **	0.032	-0.275
Physical Functioning	4	-0.300 **	0.032	-0.339	-0.301 **	0.032	-0.341
Role Functioning	2	-0.048 **	0.013	-0.054	-0.046 **	0.013	-0.052
Role Functioning	3	-0.098 **	0.014	-0.111	-0.091 **	0.013	-0.103
Role Functioning	4	-0.087 **	0.013	-0.098	-0.091**	0.013	-0.103
Social Functioning	2	-0.014	0.012	-0.016	-0.015	0.012	-0.017
Social Functioning	3	-0.031 *	0.013	-0.035	-0.031 *	0.013	-0.035
Social Functioning	4	-0.057 **	0.012	-0.064	-0.056 **	0.012	-0.064
Emotional Functioning	2	-0.024 *	0.012	-0.027	-0.024 *	0.012	-0.027
Emotional Functioning	3	-0.050 **	0.013	-0.057	-0.050 **	0.013	-0.057
Emotional Functioning	4	-0.096 **	0.012	-0.108	-0.096 **	0.012	-0.109
Pain	2	-0.034 **	0.013	-0.039	-0.034 **	0.012	-0.039
Pain	3	-0.094 **	0.013	-0.106	-0.093 **	0.013	-0.106
Pain	4	-0.111 **	0.012	-0.125	-0.110 **	0.012	-0.125
Fatigue	2	-0.035 **	0.011	-0.040	-0.035 **	0.011	-0.040
Fatigue	3	-0.055 **	0.013	-0.062	-0.054 **	0.012	-0.061
Fatigue	4	-0.061 **	0.011	-0.068	-0.060 **	0.011	-0.068
Sleep disorders	2	-0.041 **	0.011	-0.046	-0.041 **	0.011	-0.046
Sleep disorders	3	-0.051 **	0.012	-0.058	-0.051 **	0.012	-0.058
Sleep disorders	4	-0.057 **	0.010	-0.065	-0.057 **	0.010	-0.064
Lack of appetite	2	-0.042 **	0.011	-0.047	-0.041 **	0.011	-0.047
Lack of appetite	3	-0.055 **	0.013	-0.062	-0.054 **	0.013	-0.061
Lack of appetite	4	-0.074 **	0.012	-0.083	-0.073 **	0.012	-0.083
Nausea	2	-0.039 **	0.012	-0.044	-0.040 **	0.012	-0.045
Nausea	3	-0.067 **	0.013	-0.075	-0.066 **	0.013	-0.075
Nausea	4	-0.072 **	0.012	-0.081	-0.071 **	0.012	-0.080
Bowel problems	2	-0.027 *	0.012	-0.030	-0.026 *	0.012	-0.030
Bowel problems	3	-0.035 **	0.012	-0.039	-0.034 **	0.012	-0.039
Bowel problems	4	-0.049 **	0.011	-0.055	-0.049 **	0.011	-0.055

* $p<0.05$; ** $p<0.01$;

[a] Level 2= a little; Level 3 = quite a bit; Level 4 = very much; SE= Standard

3.4 Comparison of value sets from cancer patients versus the general population

Utility decrements increased monotonically with higher severity levels in most dimensions in both value sets. The cancer patient value set had more dimensions with monotonic decrements than the general population value set, requiring fewer corrections. Functional dimensions (e.g., physical, emotional) exhibited larger utility decrements than symptom dimensions in both value sets. Physical functioning had by far the biggest utility decrements in both value-sets, and pain ranked second in both. Emotional functioning ranked third in the cancer patient value set, while role functioning ranked third in the general population value set (Figure 1). For the symptoms, the patient data consistently yielded larger decrements than the general population data, except for pain. The standard errors for cancer patients' utility decrement—visualized as error bars in Figure 1—occasionally exceed these observed differences, with such error bars reflecting statistical uncertainty.

Figure 2 compares utility decrements across severity levels from the cancer patient value set and the general population value set (with the baseline set at 0), highlighting absolute differences. Cancer patients showed significantly larger decrements across all dimensions, except for certain levels of role and social functioning and sleep disorders, and all levels of pain. The largest differences were observed at level 2, particularly in symptom dimensions like fatigue, lack of appetite, and bowel problems.

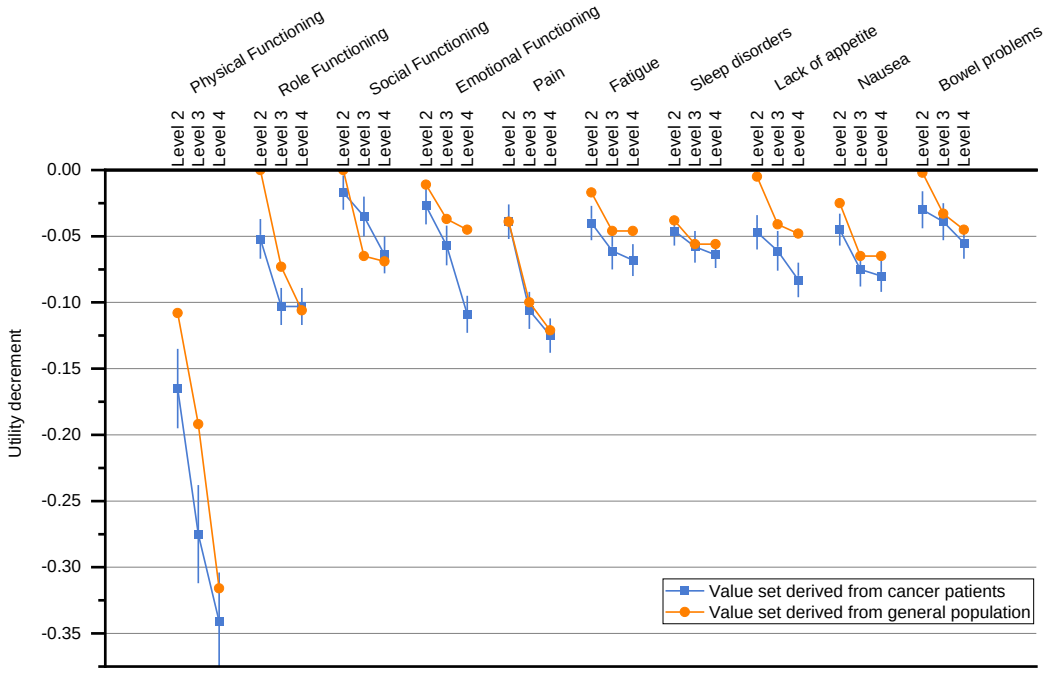

Fig. 1 QLU-C10D value sets for the cancer patient sample and the general population

Note: The value set derived from the general population is based on data from a published study[45].

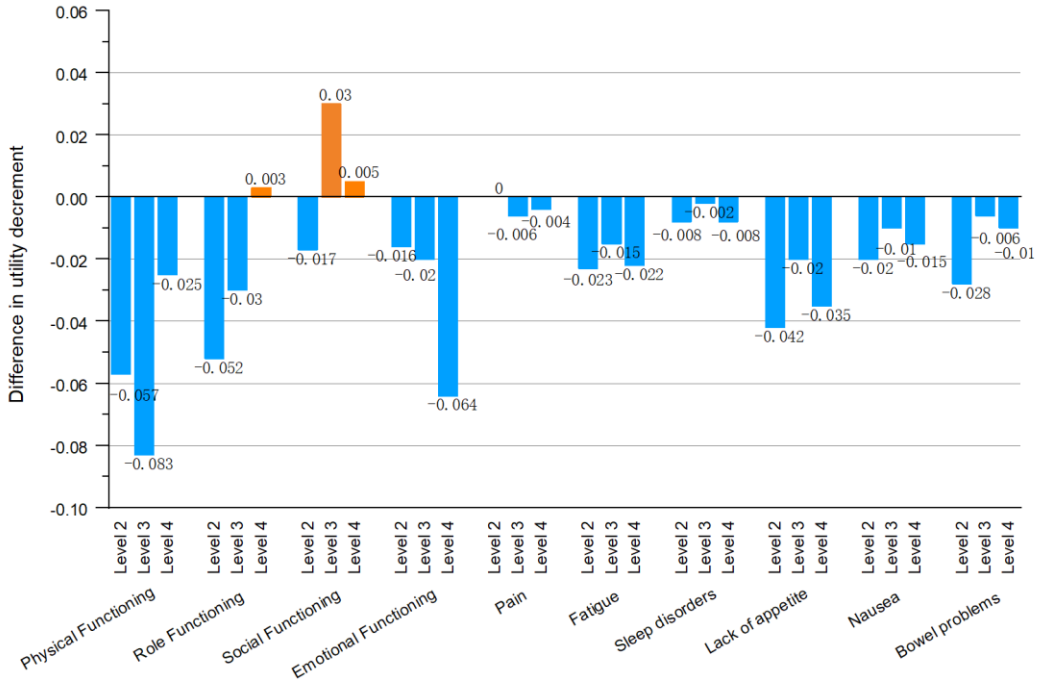

Fig. 2 Difference in value sets between cancer patients and the general population

Note: The baseline (0) represents the value set derived from the general population. The value set derived from the general population is based on data from a published study[45].

Figure 3 shows utility values for 1,000 randomly selected QLU-C10D health states, estimated using value sets from cancer patients and the general population, and ordered according to the values from the general population. Utility values from the cancer patient value set were consistently lower, especially for better health states.

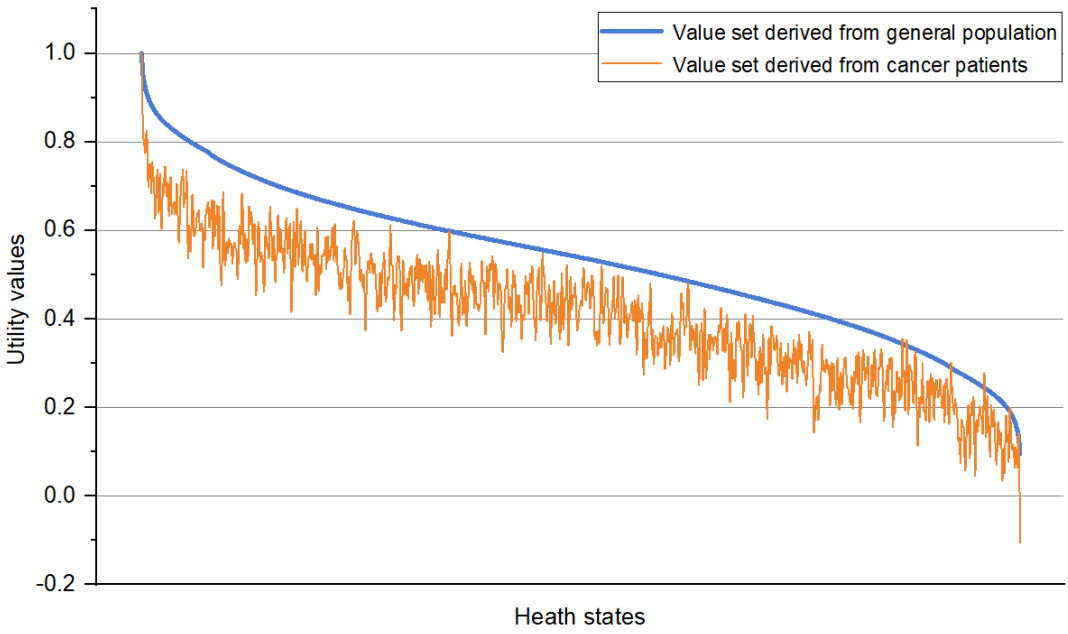

Fig. 3 QLU-C10D health states values estimated using value sets derived from Chinese cancer patients and the Chinese general population (based on data from a published study[45]), ordered according to the

values from the general population.

Figure F in the supplementary material presents utility value distributions for cancer patients using both patient and general population value sets. The cancer patient value set had a broader range (-0.035 to 1.0) and lower mean utility score (0.656) compared to the general population value set (0.091 to 1.0, mean score: 0.783), with a significant difference (P<0.01). High agreement was observed (ICC = 0.87). The Bland-Altman plot (Figure G in the supplementary material) showed proportional bias, with 95% limits of agreement ranging from -0.30 to 0.05.

3.5 Sensitivity comparison between value sets derived from cancer patients and the general population

Utility values derived from the general population and cancer patient value sets differed significantly (p<0.01) across groups stratified by cancer stage and ECOG performance status (Table 3). ES estimates comparing extreme subgroups for the cancer patient value set were larger than those for the general population for cancer stage and ECOG performance status, with larger incremental utilities.

Table 3 Sensitivity for value sets derived from Chinese cancer patients and general population

Preferences	Cancer stage	N	Mean (SD)	*p*-value	ES	Incremental utilities
Cancer patients	I	195	0.84±0.14			
	II	252	0.68±0.20			
	III	137	0.40±0.12			
	IV	25	0.26±0.23	<0.01	4.46	0.49
General population	I	195	0.90±0.09			
	II	252	0.81±0.13			
	III	137	0.60±0.13			
	IV	25	0.47±0.16	<0.01	3.48	0.33
Preferences	**ECOG**	**N**	**Mean (SD)**	***p*-value**	**ES**	**Incremental utilities**
Cancer patients	Grade 0	154	0.87±0.17			
	Grade 1	361	0.64±0.21			
	Grade 2	111	0.49±0.21			
	Grade 3	28	0.42±0.23			
	Grade 4	5	0.39±0.18	<0.01	2.76	0.48
General population	Grade 0	154	0.92±0.12			
	Grade 1	361	0.78±0.14			
	Grade 2	111	0.65±0.17			
	Grade 3	28	0.56±0.22			
	Grade 4	5	0.55±0.19	<0.01	2.47	0.37

ES: Effect size; SD: Standard deviations

Note: The analyses of cancer stage excluded leukemia cases (N=609), whereas all other analyses were conducted on the full sample (N=659).

4 Discussion

This study developed a QLU-C10D value set based on the health preferences of cancer patients in China, the country with the highest cancer incidence globally [51], and compared them with the QLU-C10D value set of the general public in China derived. This enabled us to examine differences between cancer patient versus general population value sets within the same geographic and cultural context and with similar DCE valuation methodology. The cancer patient value set—with a broader range and lower

utility values—reflects a key finding from DCE tasks: cancer patients assigned higher priority to improvements in QLU-C10D attributes than to survival time extension alone. Moreover, the cancer patient value set exhibits greater sensitivity in distinguishing between different disease subgroups, perhaps better capturing the prioritization of symptom control and quality of life improvements. These insights provide valuable guidance for selecting reference groups in health economic evaluation and underscore the importance of tailoring value sets to patient populations in specific contexts.

Among Chinese cancer patients, physical functioning, pain, and emotional functioning exhibit the highest utility decrements, which is consistent with findings from Singaporean cancer patients [50]. This similarity is likely influenced by shared Southeast Asian cultural contexts. However, in the value set for Chinese general population, emotional functioning exhibits the smallest utility decrement [45]. This discrepancy may be attributed to China's unique disease culture and traditional Chinese medicine belief that "emotions can cause cancer, and emotions can heal cancer" [68]. Utility decrements from the cancer patient value set exceed those from the general population across most dimensions, particularly symptom-related ones. The consequence is that the cancer patient health state values are generally lower than those of the general population, with the maximum difference at the pit state (-0.092 versus 0.083). The most likely explanation of preference differences that patients' have a nuanced understanding of health impairments through their lived experience and conversely that the general population tend to underestimate the impact of symptoms due to a lack of firsthand experience [9, 69]. It is important to note that cross-population differences could also reflect scale heterogeneity (i.e., variations in choice consistency) [70], as well as differences in model constraints applied to address non-monotonicities between groups. These factors may influence utility estimates beyond true preference differences, warranting cautious interpretation. While some studies suggest that patients may evaluate the impact of health impairments less than the general population [12, 13], this is often attributed to generic PBMs failing to capture patients' specific concerns about symptoms and quality of life. A systematic review found that preference differences between patients and the general population are in the same direction as in our study, and are more pronounced in disease-specific PBMs than in generic PBMs [5].

Compared to the value set of the general population, cancer patients demonstrate significantly larger utility decrements for lower severity levels (Level 2) symptoms, such as fatigue, nausea, and bowel problems, which is similar with findings from the Singaporean cancer patient value set [50]. This leads to more pronounced differences between patient versus general population value sets for better health states, while differences are smaller for worse health states. It can be inferred that cancer patients prioritize alleviating mild symptoms, likely due to their heightened sensitivity to physical vulnerability and the cumulative burden of long-term treatments. Physiologically, persistent chronic pain and treatment-related side effects significantly amplify patients' sensitivity to minor symptoms in cancer patients [71]. Psychologically, the profound anxiety about health deterioration triggered by a cancer diagnosis further intensifies this heightened sensitivity[72]. Our finding that differences in decrements between patient versus general population value sets were less pronounced at higher symptom severity levels was possibly due to patients' stronger desire for life extension and greater tolerance for discomfort [73]. This is supported by the higher time coefficient (0.882) observed in this patient cohort compared to the general population's time coefficient (0.531) [45]. A similar trend was evident in Singapore, where the time coefficient was notably higher for cancer patients (0.529) than for the general population (0.285)[50].

This study demonstrated the sensitivity of both value sets to cancer stage and performance status. However, when extrapolating intergroup differences to QALY gains based on changes in health states,

the general population value set would tend to yield lower QALY gains than the cancer patient value set due to the systematically higher utility values from the general population. In the absence of specific decision-making preferences, and in cancer-specific decision-making, the cancer patient value set may be preferred as it more accurately reflects the treatment priorities of the patients. Some authors argue that value sets from the general population may be more suitable for HTA agencies adopting a payer perspective, while patient-derived index scores are better suited for patient registries, population health studies, and personalized medicine [36]. Furthermore, the differences between this study's cancer patient value set and Singapore's highlight the need for country-specific cancer value sets. Although recruiting cancer patients and conducting face-to-face interviews are more costly, hybrid methods, such as recalibrating general population value sets from the patient perspective, could enhance QALY comparability. Incorporating patient preferences into value sets and reprioritizing items based on their perspectives [17] could also be applied to QLU-C10D preference studies, enabling flexible integration of both societal and patient-based value sets in unified evaluations. Additionally, the online elicitation of personal utility functions (OPUF) tool offers a feasible approach for constructing value sets based on the preferences of small samples of cancer patients, which warrants further exploration and validation in future research[74].

This study has notable strengths, including adherence to the internationally standardized MAUCa Consortium protocol, amended for the patient context by use of a trained interviewer to facilitate the DCE valuation task, as done in the QLU-C10D pilot patient valuation [10]. Cognitive interviews conducted in the pilot patient valuation revealed patients' need for additional effort, time and explanation to make the DCE tasks easier to understand. Our interviewers provided that extra time and explanation, resulting in more positive participant feedback on the valuation process relative to the pilot [10]. Another study strength was that it enabled comparisons of health preferences between Chinese cancer patients and general population, as well as cross-country comparisons. However, several limitations should be acknowledged. Firstly, patient recruitment was limited to two tertiary public hospitals in Harbin, China. While this approach ensured data quality and feasibility, it may limit the generalizability of our findings, as the recruited sample might not fully represent the broader cancer patient population across China in terms of socioeconomic, regional, and clinical diversity. Secondly, voluntary participation may have introduced sample bias, potentially favoring individuals with milder conditions or better treatment outcomes, while close interviewer involvement could have induced some social desirability bias. Thirdly, we did not apply education weighting (as used in the general population value set) due to inherent differences in education distributions between cancer patients and the general population; this omission may weaken the stability of extrapolating our findings for cross-group comparisons. Finally, as with all stated preference methods, our use of discrete choice experiments may be subject to hypothetical bias, where participants' stated choices may not perfectly reflect real-world decision-making.

5 Conclusion

This study developed the first QLU-C10D value set based on the preferences of Chinese cancer patients, offering a vital resource for patient-centered clinical decision-making in the country with the highest cancer incidence worldwide. Furthermore, by comparing value sets derived from cancer patients and the general population of the same country, language and culture, and using a very similar valuation protocol, the study underscores the potential differences between the two, highlighting the importance of more comprehensively integrating patient perspectives into existing evaluation frameworks.

References

1. Neumann, P.J., Goldie, S.J., Weinstein, M.C.: Preference-based measures in economic evaluation in health care. Annu Rev Public Health (2000). https://doi.org/10.1146/annurev.publhealth.21.1.587
2. Bertram, M.Y., Lauer, J.A., Stenberg, K., Edejer, T.T.T.: Methods for the Economic Evaluation of Health Care Interventions for Priority Setting in the Health System: An Update From WHO CHOICE. Int J Health Policy Manag (2021). https://doi.org/10.34172/ijhpm.2020.244
3. Cost-Effectiveness in Health and Medicine: Cost-Effectiveness in Health and Medicine. Oxford University Press (2016). https://doi.org/10.1093/acprof:oso/9780190492939.001.0001
4. Dolders, M.G., Zeegers, M.P., Groot, W., et al.: A meta-analysis demonstrates no significant differences between patient and population preferences. J Clin Epidemiol (2006). https://doi.org/10.1016/j.jclinepi.2005.07.020
5. Greiner, W., Klose, K.: Valuation of health-related quality of life and utilities in health economics. Z Evid Fortbild Qual Gesundhwes (2014). https://doi.org/10.1016/j.zefq.2014.02.004
6. Rowen, D., Azzabi Zouraq, I., Chevrou-Severac, H., van Hout, B.: International Regulations and Recommendations for Utility Data for Health Technology Assessment. PharmacoEconomics (2017). https://doi.org/10.1007/s40273-017-0544-y
7. IQWiG: https://www.iqwig.de/en/.
8. Dental and Pharmaceutical Benefits Agency (Tandvårds- och läkemedelsförmånsverket, TLV): https://www.government.se/government-agencies/dental-and-pharmaceutical-benefits-agency-tandvards--och-lakemedelsformansverket-tlv/
9. Badia, X., Herdman, M., Kind, P.: The influence of ill-health experience on the valuation of health. Pharmacoeconomics (1998). https://doi.org/10.2165/00019053-199813060-00005
10. Gamper, E.M., King, M.T., Norman, R., Loth, F.L.C., Holzner, B., Kemmler, G.: The EORTC QLU-C10D discrete choice experiment for cancer patients: a first step towards patient utility weights. J Patient Rep Outcomes (2022). https://doi.org/10.1186/s41687-022-00430-5
11. Brazier, J., Akehurst, R., Brennan, A., Dolan, P., Claxton, K., McCabe, C., Sculpher, M., Tsuchyia, A.: Should patients have a greater role in valuing health states? Appl Health Econ Health Policy (2005). https://doi.org/10.2165/00148365-200504040-00002
12. Noel, C.W., Lee, D.J., Kong, Q., Xu, W., Simpson, C., Brown, D., Gilbert, R.W., Gullane, P.J., Irish, J.C., Huang, S.H., O'Sullivan, B., Goldstein, D.P., de Almeida, J.R.: Comparison of Health State Utility Measures in Patients With Head and Neck Cancer. JAMA Otolaryngol Head Neck Surg (2015). https://doi.org/10.1001/jamaoto.2015.1314
13. Gries, K.S., Regier, D.A., Ramsey, S.D., Patrick, D.L.: Preferences for Prostate Cancer Outcomes: A Comparison of the Patient Perspective, the General Population Perspective, and a Population at Risk for Prostate Cancer. Value Health (2016). https://doi.org/10.1016/j.jval.2015.11.012
14. Peeters, Y., Stiggelbout, A.M.: Health state valuations of patients and the general public analytically compared: a meta-analytical comparison of patient and population health state utilities. Value Health (2010). https://doi.org/10.1111/j.1524-4733.2009.00610.x
15. Chai, Q., Yang, Z., Liu, X., An, D., Du, J., Ma, X., Rand, K., Wu, B., Luo, N., Schulenburg, J.M.G.V.D.: Valuation of EQ-5D-5L health states from cancer patients' perspective: a feasibility study. European Journal of Health Economics (2024).
16. Mihir, G., Ru, S., Tan, R., Ng, S., Pin, & Choo: Comparison of health state values derived from patients and individuals from the general population. Quality of Life Research An International

Journal of Quality of Life Aspects of Treatment Care & Rehabilitation (2017).
17. Ludwig, K., Ramos-Goñi, J.M., Oppe, M., Kreimeier, S., Greiner, W.: To What Extent Do Patient Preferences Differ From General Population Preferences? Value Health (2021). https://doi.org/10.1016/j.jval.2021.02.012
18. King, M.T., Costa, D.S., Aaronson, N.K., Brazier, J.E., Cella, D.F., Fayers, P.M., Grimison, P., Janda, M., Kemmler, G., Norman, R., Pickard, A.S., Rowen, D., Velikova, G., Young, T.A., Viney, R.: QLU-C10D: a health state classification system for a multi-attribute utility measure based on the EORTC QLQ-C30. Qual Life Res (2016). https://doi.org/10.1007/s11136-015-1217-y
19. Aaronson, N.K., Ahmedzai, S., Bergman, B., Bullinger, M., Cull, A., Duez, N.J., Filiberti, A., Flechtner, H., Fleishman, S.B., de Haes, J.C., et al.: The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst (1993). https://doi.org/10.1093/jnci/85.5.365
20. Giesinger, J.M., Efficace, F., Aaronson, N., Calvert, M., Kyte, D., Cottone, F., Cella, D., Gamper, E.M.: Past and Current Practice of Patient-Reported Outcome Measurement in Randomized Cancer Clinical Trials: A Systematic Review. Value Health (2021). https://doi.org/10.1016/j.jval.2020.11.004
21. Luo, N., Liu, G., Li, M., Guan, H., Jin, X., Rand-Hendriksen, K.: Estimating an EQ-5D-5L Value Set for China. Value Health (2017). https://doi.org/10.1016/j.jval.2016.11.016
22. Liu, G.G., Wu, H., Li, M., Gao, C., Luo, N.: Chinese time trade-off values for EQ-5D health states. Value Health (2014). https://doi.org/10.1016/j.jval.2014.05.007
23. Zhuo, L., Xu, L., Ye, J., Sun, S., Zhang, Y., Burstrom, K., Chen, J.: Time Trade-Off Value Set for EQ-5D-3L Based on a Nationally Representative Chinese Population Survey. Value Health (2018). https://doi.org/10.1016/j.jval.2018.04.1370
24. Liu, G.G., Guan, H., Jin, X., Zhang, H., Vortherms, S.A., Wu, H.: Rural population's preferences matter: a value set for the EQ-5D-3L health states for China's rural population. Health Qual Life Outcomes (2022). https://doi.org/10.1186/s12955-022-01917-x
25. Yang, Z., Jiang, J., Wang, P., Jin, X., Wu, J., Fang, Y., Feng, D., Xi, X., Li, S., Jing, M., Zheng, B., Huang, W., Luo, N.: Estimating an EQ-5D-Y-3L Value Set for China. Pharmacoeconomics (2022). https://doi.org/10.1007/s40273-022-01216-9
26. Wu, J., Xie, S., He, X., Chen, G., Bai, G., Feng, D., Hu, M., Jiang, J., Wang, X., Wu, H., Wu, Q., Brazier, J.E.: Valuation of SF-6Dv2 Health States in China Using Time Trade-off and Discrete-Choice Experiment with a Duration Dimension. Pharmacoeconomics (2021). https://doi.org/10.1007/s40273-020-00997-1
27. Pan, C.W., He, J.Y., Zhu, Y.B., Zhao, C.H., Luo, N., Wang, P.: Comparison of EQ-5D-5L and EORTC QLU-C10D utilities in gastric cancer patients. Eur J Health Econ (2023). https://doi.org/10.1007/s10198-022-01523-0
28. Bulamu, N.B., Vissapragada, R., Chen, G., Ratcliffe, J., Mudge, L.A., Smithers, B.M., Isenring, E.A., Smith, L., Jamieson, G.G., Watson, D.I.: Responsiveness and convergent validity of QLU-C10D and EQ-5D-3L in assessing short-term quality of life following esophagectomy. Health Qual Life Outcomes (2021). https://doi.org/10.1186/s12955-021-01867-w
29. Klapproth, C.P., Fischer, F., Rose, M., Karsten, M.M.: Health state utility differed systematically in breast cancer patients between the EORTC QLU-C10D and the PROMIS Preference Score. J Clin Epidemiol (2022). https://doi.org/10.1016/j.jclinepi.2022.09.010
30. Shaw, J.W., Bennett, B., Trigg, A., DeRosa, M., Taylor, F., Kiff, C., Ntais, D., Noon, K., King, M.T.,

Cocks, K.: A Comparison of Generic and Condition-Specific Preference-Based Measures Using Data From Nivolumab Trials: EQ-5D-3L, Mapping to the EQ-5D-5L, and European Organisation for Research and Treatment of Cancer Quality of Life Utility Measure-Core 10 Dimensions. Value Health (2021). https://doi.org/10.1016/j.jval.2021.05.022

31. Cao, Y., Li, H., Cheng, L.J., King, M.T., Kemmler, G., Cella, D., Yu, H., Huang, W., Luo, N.: A comparison of measurement properties between EORTC QLU-C10D and FACT-8D in patients with hematological malignances. Health Econ Rev (2024). https://doi.org/10.1186/s13561-024-00560-0
32. Seyringer, S., Pilz, M.J., Jansen, F., Büttner, M., King, M.T., Norman, R., Kemmler, G., Nerich, V., Holzner, B., Bottomley, A., Gamper, E.M.: Cancer-specific utility instrument for health economic evaluations: A synopsis of the EORTC QLU-C10D user manual and current validity evidence. Eur J Cancer (2025). https://doi.org/10.1016/j.ejca.2025.115235
33. King, M.T., Viney, R., Simon Pickard, A., Rowen, D., Aaronson, N.K., Brazier, J.E., Cella, D., Costa, D.S.J., Fayers, P.M., Kemmler, G., McTaggart-Cowen, H., Mercieca-Bebber, R., Peacock, S., Street, D.J., Young, T.A., Norman, R., Consortium, M.A.: Australian Utility Weights for the EORTC QLU-C10D, a Multi-Attribute Utility Instrument Derived from the Cancer-Specific Quality of Life Questionnaire, EORTC QLQ-C30. Pharmacoeconomics (2018). https://doi.org/10.1007/s40273-017-0582-5
34. Kemmler, G., Gamper, E., Nerich, V., Norman, R., Viney, R., Holzner, B., King, M., European Organisation for, R., Treatment of Cancer Quality of Life, G.: German value sets for the EORTC QLU-C10D, a cancer-specific utility instrument based on the EORTC QLQ-C30. Qual Life Res (2019). https://doi.org/10.1007/s11136-019-02283-w
35. McTaggart-Cowan, H., King, M.T., Norman, R., Costa, D.S.J., Pickard, A.S., Regier, D.A., Viney, R., Peacock, S.J.: The EORTC QLU-C10D: The Canadian Valuation Study and Algorithm to Derive Cancer-Specific Utilities From the EORTC QLQ-C30. MDM Policy Pract (2019). https://doi.org/10.1177/2381468319842532
36. Finch, A.P., Gamper, E., Norman, R., Viney, R., Holzner, B., King, M., Kemmler, G., Group, E.Q.o.L.: Estimation of an EORTC QLU-C10 Value Set for Spain Using a Discrete Choice Experiment. Pharmacoeconomics (2021). https://doi.org/10.1007/s40273-021-01058-x
37. Jansen, F., Verdonck-de Leeuw, I.M., Gamper, E., Norman, R., Holzner, B., King, M., Kemmler, G., European Organisation for, R., Treatment of Cancer Quality of Life, G.: Dutch utility weights for the EORTC cancer-specific utility instrument: the Dutch EORTC QLU-C10D. Qual Life Res (2021). https://doi.org/10.1007/s11136-021-02767-8
38. Nerich, V., Gamper, E.M., Norman, R., King, M., Holzner, B., Viney, R., Kemmler, G.: French Value-Set of the QLU-C10D, a Cancer-Specific Utility Measure Derived from the QLQ-C30. Appl Health Econ Health Policy (2021). https://doi.org/10.1007/s40258-020-00598-1
39. Revicki, D.A., King, M.T., Viney, R., Pickard, A.S., Mercieca-Bebber, R., Shaw, J.W., Muller, F., Norman, R.: United States Utility Algorithm for the EORTC QLU-C10D, a Multiattribute Utility Instrument Based on a Cancer-Specific Quality-of-Life Instrument. Medical decision making : an international journal of the Society for Medical Decision Making (2021). https://doi.org/10.1177/0272989X211003569
40. Lehmann, J., Rojas-Concha, L., Petersen, M.A., Holzner, B., Norman, R., King, M.T., Kemmler, G., Group, E.Q.o.L.: Danish value sets for the EORTC QLU-C10D utility instrument. Qual Life Res (2024). https://doi.org/10.1007/s11136-023-03569-w
41. Shiroiwa, T., King, M.T., Norman, R., Muller, F., Campbell, R., Kemmler, G., Murata, T.,

Shimozuma, K., Fukuda, T.: Japanese value set for the EORTC QLU-C10D: A multi-attribute utility instrument based on the EORTC QLQ-C30 cancer-specific quality-of-life questionnaire. Qual Life Res (2024). https://doi.org/10.1007/s11136-024-03655-7

42. Rohde, G., Lehmann, J., Pilz, M.J., Rojas-Concha, L., Holzner, B., King, M.T., Norman, R., Kemmler, G.: Norwegian and Swedish value sets for the EORTC QLU-C10D utility instrument. Quality of Life Research. https://doi.org/10.1007/s11136-024-03824-8
43. Norman, R., Mercieca-Bebber, R., Rowen, D., Brazier, J.E., Cella, D., Pickard, A.S., Street, D.J., Viney, R., Revicki, D., King, M.T.: U.K. utility weights for the EORTC QLU-C10D. Health Econ (2019). https://doi.org/10.1002/hec.3950
44. Gamper, E.M., King, M.T., Norman, R., Efficace, F., Cottone, F., Holzner, B., Kemmler, G.: EORTC QLU-C10D value sets for Austria, Italy, and Poland. Qual Life Res (2020). https://doi.org/10.1007/s11136-020-02536-z
45. Cao, Y., Xu, J., Norman, R., King, M.T., Kemmler, G., Huang, W., Luo, N.: Chinese utility weights for the EORTC cancer-specific utility instrument QLU-C10D. Qual Life Res (2024). https://doi.org/10.1007/s11136-024-03776-z
46. Jansen, F., Coupé, V.M.H., Eerenstein, S.E.J., Cnossen, I.C., van Uden-Kraan, C.F., de Bree, R., Doornaert, P., Halmos, G.B., Hardillo, J.A.U., van Hinte, G., Honings, J., Leemans, C.R., Verdonck-de Leeuw, I.M.: Cost-utility and cost-effectiveness of a guided self-help head and neck exercise program for patients treated with total laryngectomy: Results of a multi-center randomized controlled trial. Oral Oncol (2021). https://doi.org/10.1016/j.oraloncology.2021.105306
47. Yang, M., Vioix, H., Hook, E.S., Hatswell, A.J., Batteson, R.L., Gaumond, B.R., O'Brate, A., Popat, S., Paik, P.K.: Health Utility Analysis of Tepotinib in Patients With Non-Small Cell Lung Cancer Harboring MET Exon 14 Skipping. Value Health (2023). https://doi.org/10.1016/j.jval.2023.02.007
48. Kerba, M., Lourenco, R.A., Sahgal, A., Cardet, R.F., Siva, S., Ding, K., Myrehaug, S.D., Masucci, G.L., Brundage, M., Parulekar, W.R.: An Economic Analysis of SC24 in Canada: A Randomized Study of SBRT Compared With Conventional Palliative RT for Spinal Metastases. Int J Radiat Oncol Biol Phys (2024). https://doi.org/10.1016/j.ijrobp.2023.12.052
49. Byrne, A., De Abreu Lourenco, R., Govindaraj, R., Ball, D., Le, H.: A cost-effectiveness analysis of stereotactic ablative radiotherapy versus conventionally fractionated radiotherapy in the management of stage 1 non-small-cell lung cancer: Results from the TROG 09.02 CHISEL study. J Med Imaging Radiat Oncol (2024). https://doi.org/10.1111/1754-9485.13755
50. Gandhi, M., Kanesvaran, R., Rashid, M.F.B.H., Chong, D.Q., Chay, W.Y., Tan, L.Y., Norman, R., King, M.T., Luo, N.: Valuation of the EORTC Quality of Life Utility Core 10 Dimensions (QLU-C10D) in a Multi-ethnic Asian Setting: How Does Having Cancer Matter? Pharmacoeconomics (2024). https://doi.org/10.1007/s40273-024-01432-5
51. Han, B., Zheng, R., Zeng, H., Wang, S., Sun, K., Chen, R., Li, L., Wei, W., He, J.: Cancer incidence and mortality in China, 2022. J Natl Cancer Cent (2024). https://doi.org/10.1016/j.jncc.2024.01.006
52. Brown, R.C., Heines, S., Witt, A., Braehler, E., Fegert, J.M., Harsch, D., Plener, P.L.: The impact of child maltreatment on non-suicidal self-injury: data from a representative sample of the general population. BMC Psychiatry (2018). https://doi.org/10.1186/s12888-018-1754-3
53. Kessler, R.C., Andrews, G., Colpe, L.J., Hiripi, E., Mroczek, D.K., Normand, S.L., Walters, E.E., Zaslavsky, A.M.: Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med (2002). https://doi.org/10.1017/s0033291702006074
54. Rabin, R., de Charro, F.: EQ-5D: a measure of health status from the EuroQol Group. Ann Med

(2001). https://doi.org/10.3109/07853890109002087
55. Mukuria, C., Rowen, D., Mulhern, B., McDool, E., Kharroubi, S., Bjorner, J.B., Brazier, J.E.: The Short Form 6 Dimensions (SF-6D): Development and Evolution. Appl Health Econ Health Policy. https://doi.org/10.1007/s40258-024-00919-8
56. Oken, M.M., Creech, R.H., Tormey, D.C., Horton, J., Davis, T.E., McFadden, E.T., Carbone, P.P.: Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol (1982).
57. Norman, R., Viney, R., Aaronson, N.K., Brazier, J.E., Cella, D., Costa, D.S., Fayers, P.M., Kemmler, G., Peacock, S., Pickard, A.S., Rowen, D., Street, D.J., Velikova, G., Young, T.A., King, M.T.: Using a discrete choice experiment to value the QLU-C10D: feasibility and sensitivity to presentation format. Qual Life Res (2016). https://doi.org/10.1007/s11136-015-1115-3
58. Norman, R., Kemmler, G., Viney, R., Pickard, A.S., Gamper, E., Holzner, B., Nerich, V., King, M.: Order of Presentation of Dimensions Does Not Systematically Bias Utility Weights from a Discrete Choice Experiment. Value Health (2016). https://doi.org/10.1016/j.jval.2016.07.003
59. Gamper, E.M., Holzner, B., King, M.T., Norman, R., Viney, R., Nerich, V., Kemmler, G.: Test-Retest Reliability of Discrete Choice Experiment for Valuations of QLU-C10D Health States. Value Health (2018). https://doi.org/10.1016/j.jval.2017.11.012
60. Roudijk, B., Donders, A.R.T., Stalmeier, P.F.M.: Setting Dead at Zero: Applying Scale Properties to the QALY Model. Med Decis Making (2018). https://doi.org/10.1177/0272989x18765184
61. Miyamoto, J.M., Wakker, P.P., Bleichrodt, H., Peters, H.J.M.: The Zero-Condition: A Simplifying Assumption in QALY Measurement and Multiattribute Utility. Management Science (1998).
62. Norman, R., Cronin, P., Viney, R.: A pilot discrete choice experiment to explore preferences for EQ-5D-5L health states. Appl Health Econ Health Policy (2013). https://doi.org/10.1007/s40258-013-0035-z
63. Cicchetti, Domenic, V.: Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology. Psychological Assessment (1994).
64. Cohen, J.: A power primer. Psychol Bull (1992). https://doi.org/10.1037//0033-2909.112.1.155
65. Zeng, X., Sui, M., Liu, B., Yang, H., Liu, R., Tan, R.L., Xu, J., Zheng, E., Yang, J., Liu, C., Huang, W., Yu, H., Luo, N.: Measurement Properties of the EQ-5D-5L and EQ-5D-3L in Six Commonly Diagnosed Cancers. Patient (2021). https://doi.org/10.1007/s40271-020-00466-z
66. Yu, H., Zeng, X., Sui, M., Liu, R., Tan, R.L., Yang, J., Huang, W., Luo, N.: A head-to-head comparison of measurement properties of the EQ-5D-3L and EQ-5D-5L in acute myeloid leukemia patients. Qual Life Res (2021). https://doi.org/10.1007/s11136-020-02644-w
67. Pickard, D.A.S., Wilke, C.T., Lin, H.W., Lloyd, A.: Health Utilities Using the EQ-5D in Studies of Cancer. Pharmacoeconomics (2007).
68. Scheid, V.: Depression, constraint, and the liver: (Dis)assembling the treatment of emotion-related disorders in Chinese medicine. Cult Med Psychiatry (2013). https://doi.org/10.1007/s11013-012-9290-y
69. Gandhi, M., Tan, R.S., Ng, R., Choo, S.P., Chia, W.K., Toh, C.K., Lam, C., Lee, P.T., Latt, N.K.Z., Rand-Hendriksen, K., Cheung, Y.B., Luo, N.: Comparison of health state values derived from patients and individuals from the general population. Qual Life Res (2017). https://doi.org/10.1007/s11136-017-1683-5
70. Vass, C.M., Wright, S., Burton, M., Payne, K.: Scale Heterogeneity in Healthcare Discrete Choice Experiments: A Primer. Patient (2018). https://doi.org/10.1007/s40271-017-0282-4
71. Chen, X., Wang, Z., Zhou, J., Li, Q.: Intervention and coping strategies for self-perceived burden of

patients with cancer: A systematic review. Asia Pac J Oncol Nurs (2023). https://doi.org/10.1016/j.apjon.2023.100231
72. Versluis, M.A.J., van de Poll-Franse, L.V., Zijlstra, M., van Laarhoven, H.W.M., Vreugdenhil, G., Henselmans, I., Brom, L., Kuip, E.J.M., van der Linden, Y.M., Raijmakers, N.H.J.: Changes in perception of prognosis in the last year of life of patients with advanced cancer and its associated factors: Longitudinal results of the eQuiPe study. Palliat Med (2024). https://doi.org/10.1177/02692163241301220
73. Audrey, S., Abel, J., Blazeby, J.M., Falk, S., Campbell, R.: What oncologists tell patients about survival benefits of palliative chemotherapy and implications for informed consent: qualitative study. Bmj (2008). https://doi.org/10.1136/bmj.a752
74. Schneider, P., Hout, B.V., Heisen, M., Brazier, J., Devlin, N.: The Online Elicitation of Personal Utility Functions (OPUF) tool: a new method for valuing health states. Wellcome Open Research (2022).

Supplementary material

Quality of Life Utility-Core 10 Dimensions (QLU-C10D) scores based on cancer patients' health preferences are more sensitive than those based on general public's health preferences

Yiyin Cao[1]*, Haofei Li[1]*, Xu Jin[1], Madeleine T. King[2], Richard Norman[3], Ling Jie Cheng[4,5,6], Hongjuan Yu[7#], Weidong Huang[1#], Nan Luo[6]

[1] School of Health Management, Harbin Medical University, Harbin, China

[2] School of Psychology, University of Sydney, Sydney, Australia

[3] School of Public Health, Curtin University, Perth, Australia

[4] Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom

[5] Yong Loo Lin School of Medicine, National University of Singapore, Singapore

[6] Saw Swee Hock School of Public Health, National University of Singapore, Singapore

[7] Department of Hematology, The First Affiliated Hospital of Harbin Medical University, Harbin, China

*These authors have made equal contributions and shared first authorship.

[#]Correspondence:

Weidong Huang: huangweidong@hrbmu.edu.cn

Hongjuan Yu:yuhongjuan2008@163.com

Contents

Table A Health state classification system of the QLU-C10D 20

Fig. A Example of a choice set used in the discrete choice experiment valuation task to determine utility weights for the QLU-C10D 22

Fig. B An example choice set from the discrete choice experiment valuation task in Chinese as presented to survey participants 23

Fig. C Results from the QLU-C10D valuation survey feedback questions from patients 24

Data Quality Results and Discussion 25

Fig. D Distribution of time to completion by choice set **25**

Fig. E Relationship of completion time decile with model fit **26**

Fig. F Utility values of the Chinese cancer patients who participated in this study, calculated based on value sets derived from those cancer patients and the Chinese general population 26

Fig. G Bland-Altman plot of utility values of the Chinese cancer patients who participated in this study calculated based on value sets derived from those cancer patients and the Chinese general population 27

Table A Health state classification system of the QLU-C10D

Dimension	Level	Stem	Descriptor	QLQ-C30 item scores
Physical functioning	1	You have…	No trouble taking a long walk outside of the house	Item 2 (long walk) = 1
	2		No trouble taking a short walk outside of the house, but at least a little trouble taking a long walk	Item 3 (short walk) = 1 AND Item 2 ≥ 2
	3		At least a little trouble taking a short walk outside of the house, and at least a little trouble taking a long walk	Item 3 = 2 AND Item 2 ≥ 2
	4		Quite a bit or very much trouble taking a short walk outside the house	Item 3 ≥ 3 AND Item 2 ≥ 2
Role functioning	1	You are limited in pursuing your work or other daily activities…	Not at all	Item 6 = 1
	2		A little	Item 6 = 2
	3		Quite a bit	Item 6 = 3
	4		Very much	Item 6 = 4
Social functioning	1	Your physical condition or medical treatment interferes with your social or family life…	Not at all	Items 26 AND 27 = 1
	2		A little	Items 26 OR 27 = 2
	3		Quite a bit	Items 26 OR 27 = 3
	4		Very much	Items 26 OR 27 = 4
Emotional functioning	1	You feel depressed…	Not at all	Item 24 = 1
	2		A little	Item 24 = 2
	3		Quite a bit	Item 24 = 3
	4		Very much	Item 24 = 4
Pain	1	You have pain…	Not at all	Item 9 = 1
	2		A little	Item 9 = 2
	3		Quite a bit	Item 9 = 3
	4		Very much	Item 9 = 4
Fatigue	1	You feel tired…	Not at all	Item 18 = 1
	2		A little	Item 18 = 2
	3		Quite a bit	Item 18 = 3

	4		Very much	Item 18 = 4
Sleep	1	You have trouble sleeping…	Not at all	Item 11 = 1
	2		A little	Item 11 = 2
	3		Quite a bit	Item 11 = 3
	4		Very much	Item 11 = 4
Appetite	1	You lack appetite…	Not at all	Item 13 = 1
	2		A little	Item 13 = 2
	3		Quite a bit	Item 13 = 3
	4		Very much	Item 13 = 4
Nausea	1	You feel nauseated…	Not at all	Item 14 = 1
	2		A little	Item 14 = 2
	3		Quite a bit	Item 14 = 3
	4		Very much	Item 14 = 4
Bowel problems	1	You…	do not have constipation or diarrhoea at all	Items 16 AND 17 = 1
	2		have a little constipation or diarrhoea	Items 16 OR 17 = 2
	3		have constipation or diarrhoea quite a bit	Items 16 OR 17 = 3
	4		have constipation or diarrhoea very much	Items 16 OR 17 = 4

	Situation A	Situation B
In taking a long walk	You have a little trouble	You have a little trouble
In taking a short walk	You have a little trouble	You have a little trouble
You are limiting in pursuing your daily activities	Quite a bit	Quite a bit
Your physical condition interferes with your social or family life	Quite a bit	A little
You feel depressed	Quite a bit	Quite a bit
You have pain	A little	Very much
You feel tired	A little	A little
You have trouble sleeping	Not at all	Not at all
You lack appetite	Quite a bit	Quite a bit
You feel nauseated	A little	Quite a bit
You have constipation or diarrhoea	Not at all	Quite a bit
You will live in this health state for	2 years and then die	5 years and then die
Which situation would you prefer?	○	○

Fig. A Example of a choice set used in the discrete choice experiment valuation task to determine utility weights for the QLU-C10D

	健康状态 A	健康状态 B
对于长距离步行	有一点困难	有一点困难
对于短距离步行	有一点困难	有一点困难
您的工作和生活受到疾病限制	相当多	相当多
您的身体状况或医学治疗干扰到您的家庭或社交活动	相当多	有一点
您感到沮丧	相当多	相当多
您感到疼痛	有一点	非常多
您感到疲劳	有一点	有一点
您睡眠困难	没有	没有
您胃口不好	相当多	相当多
您感到恶心	有一点	相当多
您有便秘或腹泻	没有	相当多
您将在这种健康状态下生活	2 年,然后去世	1 年,然后去世
您选择哪一个？	○	○

Fig. B An example choice set from the discrete choice experiment valuation task in Chinese as presented to survey participants

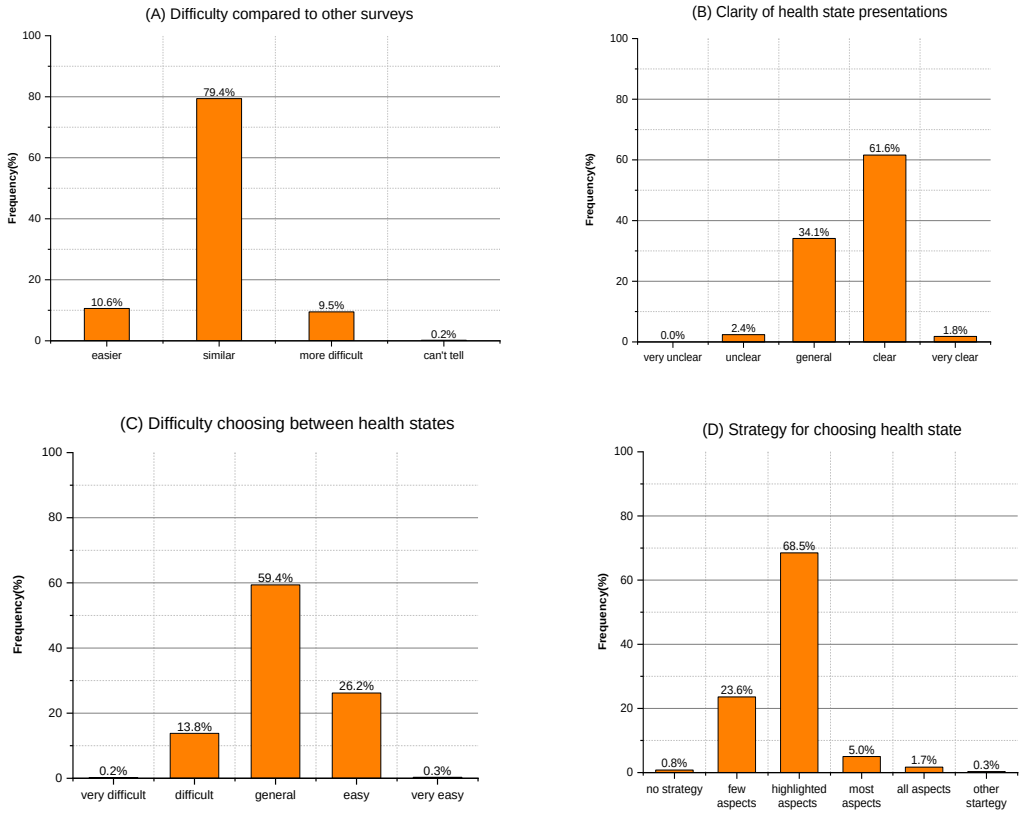

Fig. C Results from the QLU-C10D valuation survey feedback questions from patients

Data Quality Results and Discussion

Overall, as respondents became more familiar with the choice task, all respondents across the completion time deciles sped up (Figure D). Respondents in the sixth decile yielded the most statistically significant coefficients (24 of a possible 31). Although faster respondents produced fewer statistically significant coefficients, suggesting more random data, the pseudo-R^2 values of the fastest decile groups were relatively high. The slower respondents also produced fewer statistically significant coefficients, and the two slowest two deciles had the lowest pseudo-R^2 values (Figure E).

This phenomenon may be related to the following differential patterns of survey engagement in this cancer patient cohort. During the interviews, we observed that younger and healthier patients were often more willing to complete the questionnaire independently, with less need for the interviewer's explanation and interaction. They tended to complete tasks more quickly, making them more likely to appear in the earlier percentiles. In contrast, we observed that patients in poorer health or older patients generally required more time and assistance from the interviewer, even though both the interviewer's feedback and statistical results are less ideal. These informal observations were confirmed by corresponding differential composition of age and performance status across different percentiles (data not shown).

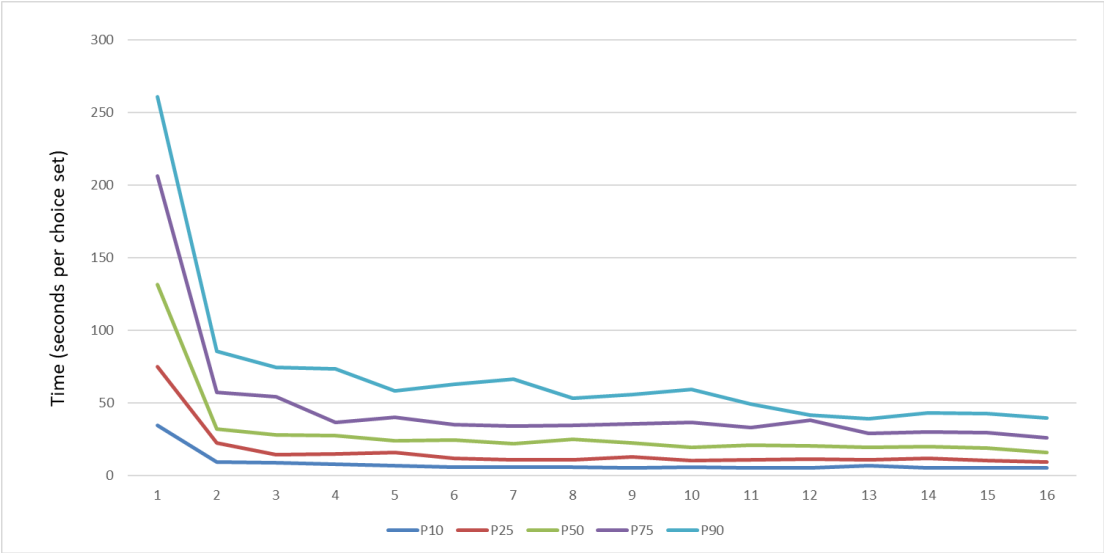

Fig. D Distribution of time to completion by choice set

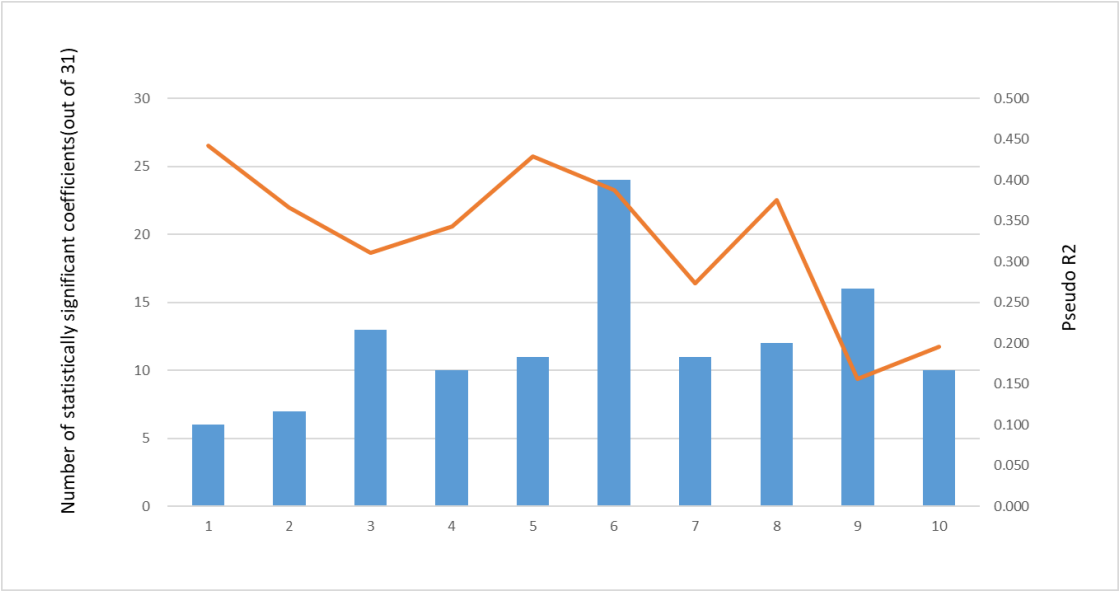

Fig. E **Relationship of completion time decile with model fit** (pseudo R-squared, orange line) and number of statistically significant coefficients (blue bars). Decile 1 contains the fastest 10% of respondents and decile 10 contains the slowest 10%.

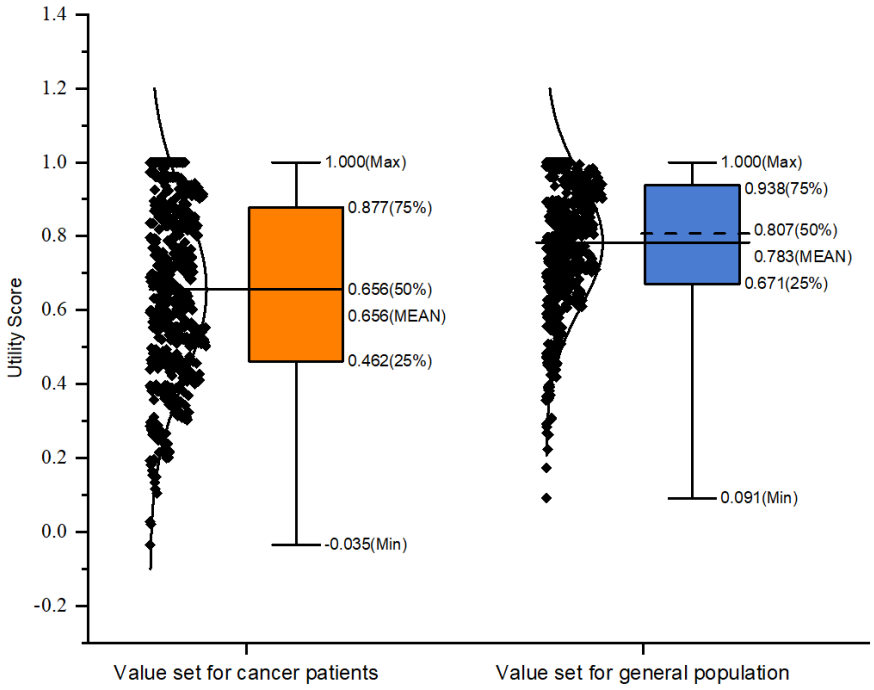

Fig. F Utility values of the Chinese cancer patients who participated in this study, calculated based on value sets derived from those cancer patients and the Chinese general population

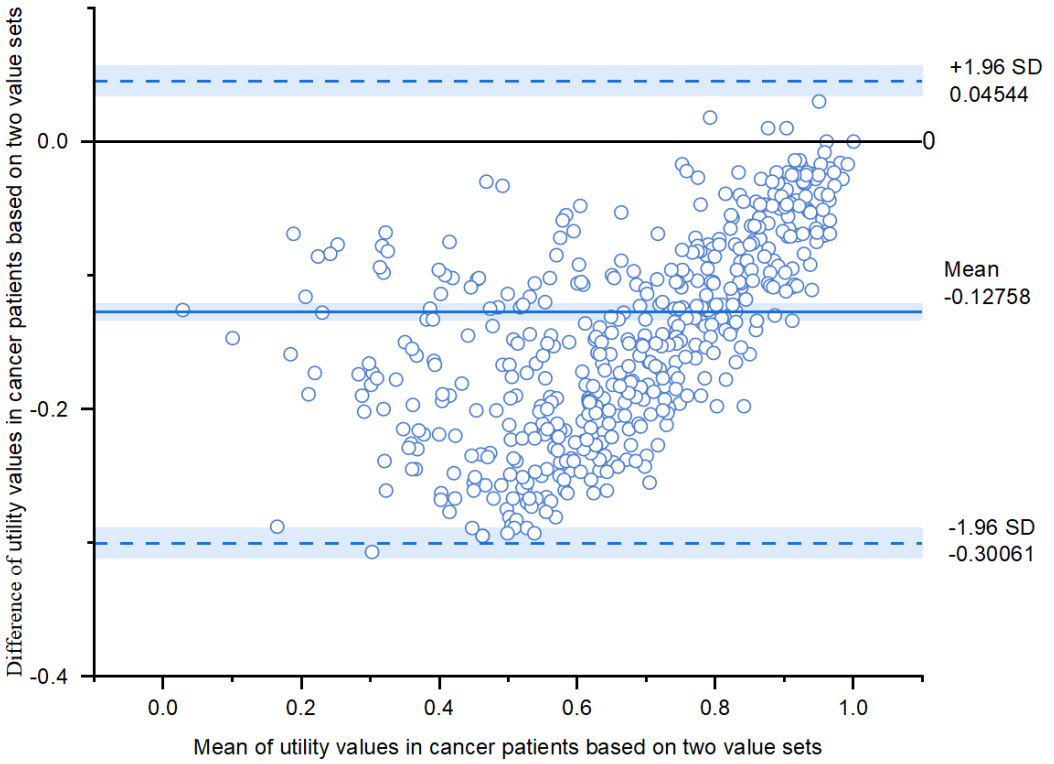

Fig. G Bland-Altman plot of utility values of the Chinese cancer patients who participated in this study calculated based on value sets derived from those cancer patients and the Chinese general population