Western Libraries Undergraduate Research Awards (WLURAs)

Volume 2024 *Vol. 2024* Article 2

2024

Do Equity-Efficiency Trade-Off Studies Assess Validity and Reliability? A Systematic Review

Victoria Chechulina
Western University

Do Equity-Efficiency Trade-Off Studies Assess Validity and Reliability? A Systematic Review

Victoria Chechulina

Supervisor: Dr. Shehzad Ali

Scholar's Electives 2200E Final Report

April 8th, 2024

3886 words

ABSTRACT

Background: Health resources are limited; therefore, policy makers must make difficult decisions about how to allocate these resources. Equity-efficiency trade-off studies use various methods to elicit individuals' preferences about the equity-efficiency trade-off in allocation decisions, with the goal of informing health policy. These methods have inherent psychometric properties, such as validity and reliability. The objective of this systematic review was to determine whether and how equity-efficiency trade-off studies assess psychometric properties, including reliability, validity, framing and cognitive effects, and robustness.

Methods: We searched MEDLINE, EMBASE, and Web of Science from inception to January 2024 for health-related equity-efficiency trade-off studies. Two reviewers independently screened titles and abstracts, followed by full-text review, data extraction, and study quality assessment.

Results: Thirty-three of the 115 included equity-efficiency trade-off studies assessed psychometric properties (28.7%, 95% CI: 20.4 - 37.0). 7.8% of included studies assessed reliability (95% CI: 2.9 - 12.7), 12.2% assessed validity (95% CI: 6.2 - 18.2), 8.7% assessed framing or cognitive effects (95% CI: 3.5 - 13.8), and 7.8% assessed robustness (95% CI: 2.9 - 12.7). Equity-efficiency trade-off studies described various methods of assessing psychometric properties, with some studies reporting high validity or reliability, and other studies reporting low validity or reliability.

Conclusions: The results of equity-efficiency trade-off studies should be interpreted critically and cautiously if there is no assessment of psychometric properties. These studies should only

be used to inform high-impact health policy decisions if researchers can demonstrate acceptable validity and reliability.

INTRODUCTION

Stated preference elicitation methods are commonly used in the health economics literature to elicit individuals' preferences regarding health states, health outcomes, or health resource allocation. Such methods typically present individuals with a hypothetical scenario where they must make trade-offs and choose between two or more alternatives.[1] Common stated preference methods include qualitative analysis, conjoint analysis, willingness to pay, and discrete choice experiments (DCEs).[2,3] Stated preference studies can reveal the value individuals associate with health states or health programs,[1] and thus have implications for health system evaluation and policy planning.[2]

In a stated preference study, individuals may be asked to make an equity-efficiency trade-off. Health resources are limited, so policy makers must balance the equity and efficiency principles when allocating these resources.[4] The efficiency principle states that overall health should be maximized, and that more health in the population is beneficial, irrespective of how health is distributed.[5] Conversely, the equity principle postulates that some health gains are more valuable than others if they minimize unfairness in the distribution of population health.[6] Many studies have investigated the equity-efficiency trade-off to ascertain whether individuals prefer health resources to be allocated efficiently or equitably, with the goal of informing health policy. Systematic reviews of these studies have summarized individuals' preferences in the equity-efficiency trade-off.[7–9]

Psychometric properties describe the validity and reliability of measurement instruments.[10] Validity is the degree to which the instrument measures exactly what it proposes to, while reliability is the ability of an instrument to produce consistent results over time and space.[11] Specific examples of psychometric properties include test-retest reliability, criterion validity, content validity, and construct validity.[12] There exist many techniques to assess the validity and reliability of DCEs and other stated preference methods. Since DCEs present hypothetical scenarios, the validity of DCEs can be evaluated by comparing the stated preferences elicited by the experiment with the revealed preferences made by individuals in real life.[13] Similarly, the validity of DCEs can be assessed by determining the extent to which all components of the DCE survey have prompted participants to make choices that are in line with their true preferences.[12] Reliability of a choice experiment can be assessed by providing the same choice sets multiple times to the same individual, or different individuals, and identifying whether the same result has been produced.[12] Ultimately, it is crucial to assess validity and reliability of stated preference methods if the results of these studies are to be used to inform health policy.[14] Assessment of psychometric properties lends credibility to the method used.[15]

While several systematic reviews of equity-efficiency trade-off studies have been published,[7–9,16,17] these reviews do not discuss psychometric properties as they relate to the included studies. To our knowledge, there have been no systematic reviews published to date that examine the psychometric properties of equity-efficiency trade-off studies specifically. It is unknown whether and how equity-efficiency trade-off studies assess the psychometric properties of their measurement instruments. This gap in the literature has important implications for policy development. If the results of equity-efficiency trade-off studies are to

be used in health resource allocation decisions, researchers should assess the psychometric properties of their experiments to ensure meaningful and accurate results.

The objective of this systematic review was to determine whether health-related equity-efficiency trade-off studies assess the psychometric properties of their measurement instruments. We hypothesized that most health-related equity-efficiency trade-off studies would not assess or discuss the psychometric properties of their measurement instruments. This review fills a gap in the literature by quantifying the number of equity-efficiency trade-off studies that discuss psychometric properties, an important consideration if these studies are to be used in health policy decisions.

METHODS

This systematic review was designed and reported according to the Preferred Reporting Items for Systematic reviews and Meta-analyses (PRISMA) 2020 guidelines.[18]

Eligibility Criteria

Study inclusion criteria were determined prior to study screening. We included studies that were: (1) peer-reviewed; (2) in English; (3) elicited value judgements and preferences from individuals with respect to health allocation decisions or asked individuals to make a health-related equity-efficiency trade-off; and (5) included equity or equality as a dimension of choice. Any stated preference study was eligible for inclusion regardless of the preference elicitation technique, such as DCEs, conjoint analyses, or best-worst scaling methods. Theoretical studies, studies unrelated to health, studies that did not include equity or equality as a dimension of choice, and studies that did not present survey results were excluded. There were no

restrictions placed on year of publication, study sample, or study setting. Peer-reviewed original research was included, and conference abstracts were excluded.

Information Sources

Three databases were searched on December 4th, 2023: Ovid MEDLINE (1946 – 04/12/2023), Ovid EMBASE (1947 – 04/12/2023), and Web of Science (1900 – 04/12/2023). Searches were updated on January 12th, 2024. Reference lists of systematic reviews of equity-efficiency trade-off studies were manually searched to identify potentially eligible studies.[7–9]

Search Strategy

The search strategy was designed to capture health-related equity-efficiency trade-off studies by combining three concepts: *health* AND *preference elicitation study* AND *equity or equality*. Terms for each of these three concepts were identified by consulting previous systematic reviews of equity-efficiency trade-off studies.[7–9] No limits were placed on date or language. The sensitivity of the search strategy was validated by ensuring that studies included in previously published systematic reviews were captured by the search strategy. The full search strategies can be found in Appendix A.

Study Selection

Records identified from the searches were uploaded to Covidence for screening. Two reviewers independently screened all titles and abstracts, followed by full-text review. Conflicts were resolved by a third reviewer. Two reviewers independently extracted data from each included study. Discrepancies were discussed with a third reviewer. The results presented in

this version of the manuscript are from single data extraction, while duplicate data extraction is still ongoing.

Data Items and Outcomes

The primary outcome for this review was the proportion of equity-efficiency trade-off studies that assessed, discussed, or evaluated the psychometric properties of their measurement instruments. For each study included, reviewers identified whether the study assessed reliability, validity, framing and cognitive effects, or robustness. Although framing effects, cognitive effects, and robustness are not traditional psychometric properties, data were extracted with respect to these effects as well because such effects have been discussed in the equity-efficiency trade-off literature.[19] If a study assessed psychometric properties, reviewers identified how these properties were assessed, and what the findings were. If a study provided quantitative values for the psychometric properties of its measurement instrument, these values were extracted. If a study discussed psychometric properties narratively, this information was synthesized.

Data were collected for the following study characteristics in studies that assessed psychometric properties: stated preference elicitation method; sample size; geographic location; type of sample (e.g., general public, policymakers, healthcare professionals, etc.); and survey administration method (e.g., face-to-face interview, online survey, etc.).

Study Quality Assessment

Study quality was assessed using the PREFS checklist, which is suitable for assessing data quality across different preference elicitation methods.[20] The checklist assesses five study

domains: purpose; respondent sampling; explanation of methods; findings; and significance testing. Each item is scored 0 (unacceptable) or 1 (acceptable) and each study receives a score from 0 to 5.[21] Study quality assessment was performed for studies that assessed psychometric properties independently by two reviewers, and discrepancies were resolved with a third reviewer. At the time of writing this manuscript, study quality assessment is ongoing.

Data Synthesis

We calculated the proportion of included studies that assessed any psychometric property, and the proportion of included studies that assessed each of reliability, validity, framing or cognitive effects, and robustness, with a 95% confidence interval for the proportion. For each included study that assessed psychometric properties, we summarized the findings of the assessment in a table. Narrative synthesis was used to describe the reporting of psychometric properties in the equity-efficiency trade-off literature.

Confidence in cumulative evidence was not formally assessed using GRADE because other systematic reviews of equity-efficiency trade-off studies have not done so in the past. Due to the wide heterogeneity of study methods, a meta-analysis was not appropriate, and review results were reported descriptively.

RESULTS

The database searches returned 5279 results (1273 MEDLINE, 1424 EMBASE, 2582 Web of Science) and hand searching of reference lists returned three additional citations. After deduplication, 3240 titles and abstracts were screened. The full texts of 200 articles were reviewed, and 115 studies were included. The most common reason for exclusion was the study

not being a preference elicitation study (n = 39). The study identification, screening, and inclusion process is summarized in a PRISMA flow diagram (Figure 1).[18]

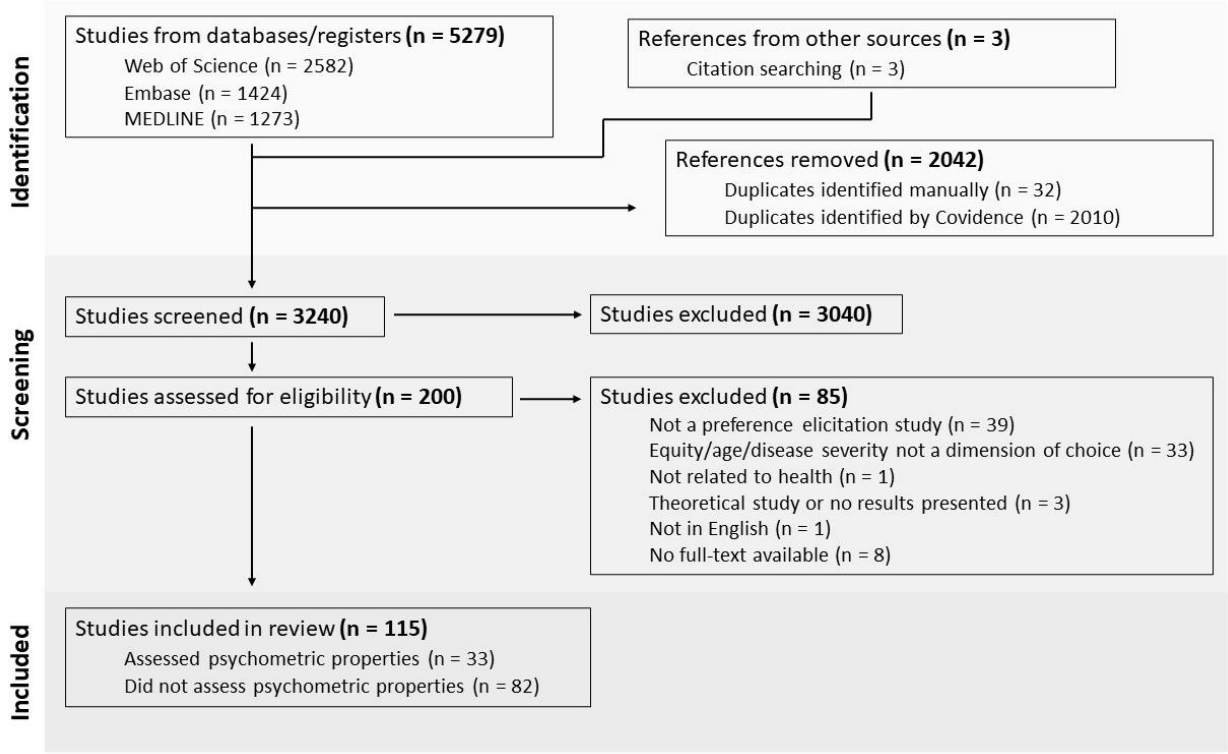

Fig. 1. PRISMA flow diagram.

Study design characteristics of the studies that assessed psychometric properties are summarized in Table 1. Of the 115 included studies, 33 (28.7%, 95% CI: 20.4 – 37.0) assessed psychometric properties.[22–54] The other 82 included studies did not assess or discuss psychometric properties.[55–136]

Table 1. Study design characteristics of studies that assessed psychometric properties.

Study	Preference elicitation method	Sample size	Geographic location	Type of sample	Administration format
Gulácsi 2012	Discrete choice experiment	200	Hungary	National sample	Face-to-face interviews and online questionnaires
Attema 2022	Person trade-off	500	The Netherlands	National sample	Online
Stolk 2005	Priority ranking experiment	65	The Netherlands	Students, researchers, health policy makers	Not stated
Ratcliffe 2000	Conjoint analysis	303	Britain	Local sample of university employees	Questionnaire sent by mail
Werner 2009	Survey	624	North Israel	Local sample of community-dwelling adults	Face-to-face interview
Luyten 2019	Discrete choice experiment	750	Belgium	National sample	Survey (not stated whether online or in-person)
Damschroder 2004	Person trade-off	95	Philadelphia, USA	Local sample (prospective jurors)	Computer-administered survey and face-to-face interview
Jehu-Appiah 2008	Discrete choice experiment	37	Ghana	Specialized group of directors	In-person
Reckers-Droog 2019	Person trade-off tasks and choice tasks	1025	The Netherlands	National sample	Online questionnaire
Reckers-Droog 2021	Willingness-to-pay	2023	The Netherlands	National sample	Online
Ubel 1996	Choice experiment	169	Pennsylvania, US	Local sample of jurors	Paper questionnaire
Whitty 2014	Discrete choice experiment and best-worst scaling	930	Queensland, Australia	Local sample	Online
Lancsar 2011	Discrete choice experiment	587	England	National sample of the general population	Face-to-face using Computer Assisted Personal Interview (CAPI)
Petrou 2013	Person trade-off	2500	United Kingdom	National sample	Online survey
Aidem 2017	Focus groups and semi-structured interviews	27	Norway	Hospital administrators, policymakers, practitioners, seniors, and university students	In-person
Whitty 2015	Discrete choice experiment	1994	Queensland and South Australia	Local sample	Online survey
Cookson 2018	Questionnaire with pairwise choices	60	York, England	Local sample	Face-to-face

Study	Preference elicitation method	Sample size	Geographic location	Type of sample	Administration format
Robson 2017	Questionnaire with pairwise choices	244	England	General public	Online
Rowen 2016	Discrete choice experiment	3669	United Kingdom	General population	Online
Ahlert 2017	Questionnaire	166	Germany	Local sample of university students	In-person
Ali 2017	Choice experiment	135	York, England	Local sample	Online and face-to-face
Li 2022	Questionnaire	1862	United States	Physicians, compared with the general public, an 'elite' subsample of Americans, and a nationwide sample of medical students	Online survey
Ubel 2001	Questionnaire	615	Philadelphia, USA	Local sample of randomly selected jurors	In-person survey
Ubel 1999	Choice experiment	479	Philadelphia, US	Local sample of jurors	In-person survey
Baker 2010	Qualitative questions, discrete choice experiment, matching, standard gamble	587 for the relativities study, 409 for the valuation study	England	National sample	Computer-assisted personal interview
Abásolo 2013	Choice experiment	1211	Spain	National sample	Face-to-face interview
Schoon 2022	Integrated citizens jury and discrete choice experiment	27	Taiwan	Local sample	Face-to-face
McKie 2019	Small-group discussion and interview	66	Victoria, Australia	Local sample	Semi-structured, face-to-face, small-group discussion
Comerford 2023	Experiment with pairwise choices	495	United States and United Kingdom	National sample	Online survey
Baltussen 2006	Discrete choice experiment	30	Ghana	Health policy makers	In-person survey
Ubel 1996	Choice experiment	169	Pennsylvania, US	Prospective jurors	Face-to-face survey
Green 2009	Choice experiment	261	Southampton, England	Local sample	Face-to-face interview
van Exel 2015	Q methodology	294	Ten countries: Hungary, Palestine, Poland, Denmark, Norway, France, Sweden, the UK, the Netherlands, Spain	General public, national samples	

9 studies assessed reliability (7.8%, 95% CI: 2.9 - 12.7), 14 assessed validity (12.2%, 95% CI: 6.2 - 18.2), 10 assessed framing or cognitive effects (8.7%, 95% CI: 3.5 - 13.8), and 9 assessed robustness (7.8%, 95% CI: 2.9 - 12.7). Some studies assessed more than one psychometric property. These results are summarized in Figure 2.

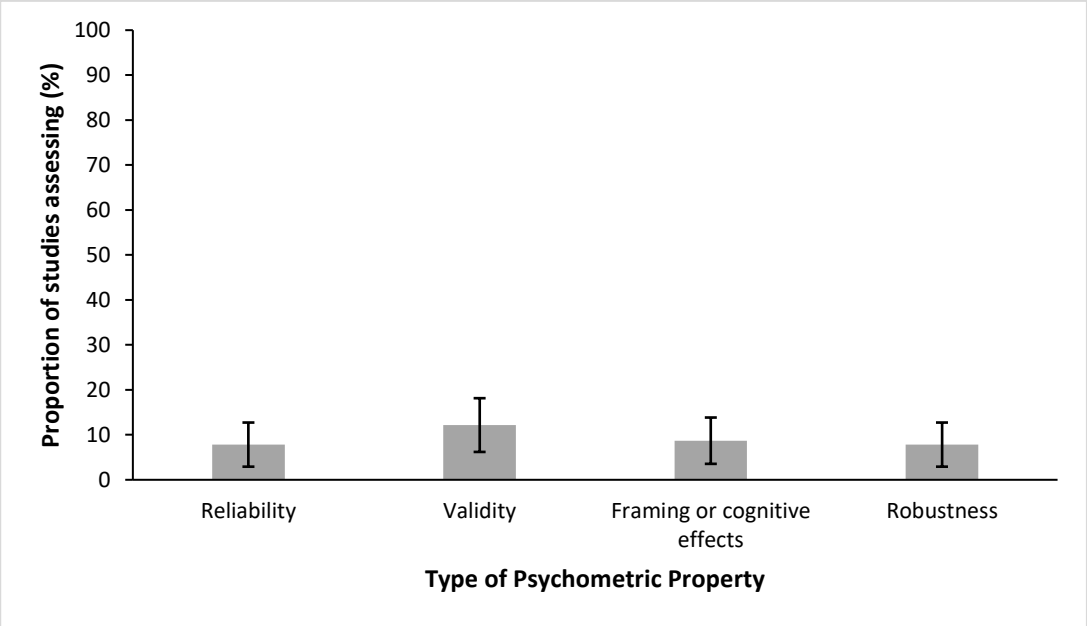

Fig. 2. Percentage of equity-efficiency trade-off studies assessing psychometric properties (n=115). Error bars show 95% confidence interval for the proportion.

Tables 2, 3, 4, and 5 present the characteristics of psychometric properties in studies that assessed reliability, validity, framing or cognitive effects, and robustness, respectively.

Of the nine studies that assessed reliability, five assessed internal consistency or internal reliability[31,38,42,45,51] and four assessed test-retest reliability.[22,24,46,50] Cronbach's alpha was used as a measure of internal consistency in three studies, with two studies reporting good internal consistency[38,42] and one study reporting modest internal consistency.[45] To assess test-retest reliability, researchers provided repeated choice tasks[24] or a repeated question at the end of

the questionnaire.[50] In one study, a subsample of individuals was asked to repeat the survey one week after the initial survey,[22] and in another study, investigators administered the survey both online and in-person two weeks apart and compared the results from both modes of administration.[46] Generally, researchers reported high test-retest reliability. For example, 81.1% of respondents made the same choices when provided with a repeated question,[50] and in another study, there was strong or excellent agreement between the values provided at the initial test and at retest.[22]

Fourteen studies assessed validity. Some investigators did not specify the type of validity they assessed, while others specified face validity[40] or theoretical validity.[47,54] Several studies ascertained the validity of their measurement instruments by comparing the results with *a priori* expectations or with previously published literature or policy.[27,33,41,44,47,52,54] Most of these studies reported results consistent with *a priori* expectations and published literature, except for one study that reported low internal validity because respondents did not match rational behaviour expectations.[41] In one study, participants were given the option to 'opt out' when faced with a difficult allocation decision.[36] Very few participants decided to opt out, suggesting that they were not avoiding difficult decisions and that their stated preferences reflected their true beliefs, indicating good validity.[36] Investigators in another study reported high validity because participants' choices on the questionnaire were highly consistent with the verbal statements they made to explain their reasoning.[23] Comerford and colleagues assessed validity by comparing individuals' preferences in an abstract scenario with their preferences in a real-world COVID-19 pandemic scenario. In the real-world scenario, participants were less likely to 'level down' and choose as if an additional year of life has negative utility if offered to the most

privileged.[45] This discrepancy in preferences between real-world and abstract resource allocation scenarios suggests that abstract scenarios may not be valid in eliciting individuals' true preferences. Overall, most equity-efficiency studies reported acceptable validity of their measurement instruments.

Ten studies assessed framing or cognitive effects. Half of these studies varied question or attribute framing,[22,27,34,35,39] and others assessed order effects by presenting questions in different orders.[37,48] Two studies used a pilot study or focus groups to rule out framing and order effects prior to the main investigation.[26,31] Five studies did not observe any significant framing or order effects.[22,31,34,37,48] Conversely, three studies found that respondents' choices were affected by question framing or order.[27,35,39] In one study, investigators attempted to reduce pro-egalitarian cognitive bias by randomizing participants to receive an e-learning intervention designed to prompt participants to think about questions in a more complete manner prior to taking the questionnaire.[28] Respondents that received the e-learning intervention still displayed high levels of inequality aversion, but there was a substantial new minority of non-egalitarian responses.[28] In another study, Ali and colleagues tested four different cognitive biases based on question framing. They found that presenting small vs. unrealistically large inequality reductions in their hypothetical scenarios, and population- vs. individual-level descriptions did not affect health inequality aversion.[19] The presentation of concrete scenarios instead of abstract scenarios, and an online administration format led to some reductions in health inequality aversion, but the median respondent was still inequality averse.[34] In summary, while some studies reported that there were no major framing effects present in their study, others reported that framing effects significantly affected respondents'

choices, and others reported only a modest framing effect. As such, it is unclear to what extent framing effects influence the results of equity-efficiency trade-off studies.

Nine studies assessed robustness. Several studies reported sensitivity analyses as a method of checking robustness.[30,32,54] Three studies assessed robustness by repeating their analyses after excluding a certain group of participants. One study excluded participants who showed no real preferences in the choice task.[32] Investigators in this study reported that their results were affected by excluding participants who showed no real preference.[32] Another study repeated analyses after excluding participants who may not have understood or engaged with the survey, found the survey difficult, took a long or short time to complete the survey, or selected the same response for all questions.[29] This study found that excluding certain participants affected the magnitude of all calculated coefficients, but this effect was only statistically significant for a small number of models and the impact was not systematic.[29] A third study excluded participants who reported a lower clarity or certainty score for the choice task, took a long or short time to complete the task, or stated a lower willingness-to-pay for larger quality of life gains.[54] Other equity-efficiency trade-off studies assessed robustness by including individuals with inconsistent responses[24] and control participants.[33,45] These studies found that results were robust to the inclusion of inconsistent responders and controls. One study assessed whether median inequality aversion parameters were sensitive to larger sample sizes, and found that these parameters were indeed robust to sample size.[30] In general, most studies reported that their results were robust to the inclusion or exclusion of certain participants, variations in methods, and differences in sample size.

Table 2. Characteristics of psychometric properties in studies that assessed *reliability.*

Study	Type of psychometric property	How was it assessed?	Findings
Comerford 2023	Inter-item reliability	Cronbach's alpha was used to assess inter-item reliability of the sacred values scale, where 'sacred values' were a covariate and potentially influenced respondents' health inequality aversion.	The sacred values scaled had only modest inter-item reliability. Cronbach's alpha = 0.556, and a desirable value would be 0.7.
Stolk 2005	Internal consistency	An internal consistency check to determine the robustness of the preferences elicited in the paired comparison task.	The observed proportions agreed well with the expected proportions. The coefficient of consistency showed a mean value of 0.947, indicating a high level of consistency.
Werner 2009	Internal reliability	Cronbach's alpha was calculated for each of the three value indexes that were constructed to reflect participants' value orientations.	The first index had excellent internal reliability (alpha = 0.84), the second had very good internal reliability (alpha = 0.76), the third had good internal reliability (alpha = 0.74).
Luyten 2019	Internal consistency	Cronbach's alpha was calculated for the life orientation, optimism, and pessimism variables. The association between these dispositions and priority setting was then calculated.	The internal consistency of the three variables was acceptable [Cronbach's α = 0.75 (life orientation), 0.78 (pessimism) and 0.72 (optimism)].
Damschroder 2004	Internal consistency	Three measures of internal consistency were calculated: ordinal consistency, reference group size consistency, and triad consistency.	Only 5% of responses violated criteria for ordinality, and the odds of having violations in ordinality did not differ between the computer and face-to-face conditions. Both groups displayed significant bias with respect to reference group size consistency. The two elicitation modes did not differ in triad inconsistency.
Whitty 2014	Test-retest reliability	Repeated choice tasks were performed for both the Discrete Choice and Profile Case Best-Worst Scaling Methods.	Consistent responses were given to the DCE repeat choice task by 75.7% of respondents. Consistent responses were given to the BWS repeat choice task by 64.5% for the most preferred attribute level, 49.4% for the least preferred attribute level, and 35.5% for both the most and least preferred.
Petrou 2013	Test-retest reliability	The first 500 individuals (one fifth of the sample) were asked to complete a re-test survey with identical questions, in the same order and using the same online format, approximately one week after the first survey. These individuals did not have access to their previous responses.	There was strong agreement (0.7-0.8) between values at test and retest for 3 of 8 person trade-off questions, and excellent agreement (>0.8) for 5 of 8 person trade-off questions. Reliability was high.
Attema 2022	Test-retest reliability	A repeated question was presented at the end of the questionnaire to see if participants would choose the same option.	81.1% of the subjects made the same choice in the repeated questions, suggesting that reliability was fairly high.
van Exel 2015	Test-retest reliability	Materials were designed such that questions could be administered in person or online. A repeated in-person and online interview pilot study was conducted two weeks apart.	"Comparability of results from both modes of administration was good (ρ=0.8)."

Table 3. Characteristics of psychometric properties in studies that assessed *validity.*

Study	Type of psychometric property	How was it assessed?	Findings
Baltussen 2006	Validity	The signs of relations were examined, and the results were compared to the priorities of the international health community.	All relations had the expected sign and the results showed a large overlap with the priorities of the international health community. These observations support the face validity of the results.
Ahlert 2017	Validity	Researchers compared individual's choices in the questionnaire with verbal statements meant to elucidate their reasoning.	Individuals' verbal statements were highly consistent with their choices on the questionnaire.
Damschroder 2004	Validity	A computer-based administration mode was compared to a face-to-face interview format.	The computerized protocol produced results of similar quality to the face-to-face protocol.
Jehu-Appiah 2008	Validity	Rankings were compared to the rankings produced from a simple DCE conducted on the same sample.	The results showed a strong correlation with the simple exercise (Spearman rank order correlation = 0.79), suggesting that validity was high.
Reckers-Droog 2021	Theoretical validity	Investigators hypothesized that willingness-to-pay would be higher for larger-sized quality of life gains and for respondents with a higher household income. Results in support of this hypothesis would indicate acceptable theoretical validity.	Higher household income of the respondent and higher quality-of-life gains corresponded to higher willingness to pay, suggesting acceptable theoretical validity.
Ubel 1996	External validity	The results of the study were compared with current organ allocation policies to determine if these policies are valid.	"The validity of basing transplant policies, at least in part, on public values is supported by the similarities between present allocation policies and the results of this study."
Aidem 2017	Validity	“To address validity, the author interviewed stakeholders from different levels of the health system about the same questions of interest, documented all research activities for critical evaluation of the methodology and shared findings with a sample of respondents to assess the authenticity of the author’s interpretations.”	Validity was adequate.
Whitty 2015	Face validity	"Extensive pilot testing was undertaken to confirm the face validity of the instrument, prior to main data collection."	Face validity was adequate.
Li 2022	External validity	Investigators conducted earlier studies and consulted other literature to establish the external validity of their experimental method.	The experimental method used by the investigators was externally valid.
Gulácsi 2012	Internal validity	“Investigated through *a priori* expectations based on rational behavior that maximizes health benefits.”	Respondents did not match the rational behaviour set from the *a priori* expectations, not valid.
Schoon 2022	Theoretical validity	Signs of coefficients in the model were calculated and compared with *a priori* expectations.	The signs of coefficients in the model were consistent with expectations when *a priori* expectations were applied, so the model was theoretically valid. Post-citizens jury DCE is a valid method of eliciting preferences.

Study	Type of psychometric property	How was it assessed?	Findings
Comerford 2023	Validity	Two surveys were given - one describing an abstract allocation scenario similar to previous equity-efficiency trade-off studies, and one describing real-world COVID-19 conditions. Meant to examine whether 'abstract' preferences match those that people would make in real life.	In the real-world COVID-19 scenario, participants are less likely to 'level down' - i.e. choosing as if an additional year of life has negative utility if offered to the most privileged. This suggests that in real-world scenarios, participants may prefer efficiency over equity, in contrast with previous equity-efficiency studies which commonly show a strong health inequality aversion in abstract scenarios.
Ratcliffe 2000	Validity	Results from the regression model were used to evaluate a priori expectations. It was expected that individuals would allocate more donor livers to a group if they had a longer expected survival, and less livers if they had alcoholism, were older, or were a re-transplant.	The results provide some evidence for the model's validity.
Green 2009	Internal validity	Respondents were given an option to 'opt out' when faced with a difficult decision of allocating healthcare resources.	Very few (5%) respondents chose to opt out, suggesting that the most common preference of dividing resources equally may be a true preference, rather than participants avoiding difficult decisions.

Table 4. Characteristics of psychometric properties in studies that assessed *framing and cognitive effects*.

Study	Type of psychometric property	How was it assessed?	Findings
Lancsar 2011	Question formats, graphical presentation of attributes	Several focus groups were conducted to determine the effect of question framing.	Focus groups were necessary to narrow down the list of attributes with a clear definition of each and determine the best framing for each question.
Petrou 2013	Attribute framing effect	"To assess whether estimated age-related weights for health gains were influenced by the framing of the trade-off questions, the reference number of individuals in each age group was set at 100 for one half of the study participants and at 1000 for the other half of study participants."	Adopting alternative referents in the trade-off questions did not have a significant effect on study results.
Cookson 2018	"Pro-strict egalitarian" cognitive bias	Participants were randomized to receive an e-learning intervention prior to taking the questionnaire to allow the respondents to think about questions in a more complete manner.	Respondents that received the e-learning intervention still had a high level of health inequality aversion but there was a substantial new minority of non-egalitarian responses.
Ali 2017	Cognitive biases based on question framing	Investigators tested four different cognitive biases based on question framing and presentation: (1) small vs. unrealistically large health inequality reductions (2) population-level vs. individual-level descriptions of health inequality reduction (3) concrete vs. abstract intervention scenarios (4) online vs. face-to-face mode of administration.	Small vs. unrealistically large inequality reductions and population- vs. individual-level descriptions did not affect health inequality aversion. Concrete scenarios and online format led to some reductions in health inequality aversion but the median respondent was still health inequality averse.
Ubel 2001	Framing effects	Investigators varied the attributes of scenarios and the order in which scenarios were presented	Individuals' preferences varied significantly depending on the ways in which scenarios were framed. Equity-efficiency trade-off preferences are susceptible to framing effects.
Ubel 1999	Framing effects - question wording	Minor changes were made to question wording - by clarifying the scenario, prompting subjects to think about self-interest, and giving subjects the explicit vs. no explicit option to divide resources evenly.	Less subjects prioritized severely ill patients when they were given the explicit option to divide resources evenly. Minor changes in wording do affect people's preferences for prioritizing severely ill patients.
Abásolo 2013	Titration effects (based on the order in which the questions are presented, and the 'gap' between the options)	Some participants were presented with random question order, others were presented with a 'titrated' question order, with gradually increasing differences between each and every question.	No significant framing effect was found - it doesn't seem that the titration effect influences participants' choices.
Stolk 2005	Framing effects based on presentation of health states	A pilot study was conducted.	Rank order effects were not affected by the presentation of health states or labels in the pilot study - no framing effects were observed.
McKie 2019	Order effects	Health states were presented in different orders. Participants were asked to partake in a process of critical self-examination that was designed to prompt them to think beyond their initial reactions, with the goal of eliminating order effects.	Unlike previous studies, this study did not observe order effects - order effects were eliminated.
Ubel 1996	Framing effects	Investigators tested whether framing recipients of transplant organs as distinct groups vs. individual recipients affected people's allocation preferences.	Subjects were more inclined to use prognostic information in determining their choices when patients were presented as individuals rather than groups.

Table 5. Characteristics of psychometric properties in studies that assessed *robustness.*

Study	Type of psychometric property	How was it assessed?	Findings
Reckers-Droog 2019	Robustness	Sensitivity analyses: repeating the analyses excluding respondents who reported a low score.	Robustness checks indicated that the obtained ratios were pulled downwards by including respondents with no real preference.
Whitty 2014	Model robustness analysis	By examining the performance of these models when including responders with inconsistent responses, the study evaluates the resilience and stability of the models to variations in respondent behavior,	“Models were robust to the inclusion of inconsistent responders.”
Robson 2017	Sensitivity analysis	Sensitivity analyses were conducted to determine whether median inequality aversion parameters were sensitive to larger samples.	Median inequality aversion parameters were robust.
Rowen 2016	Robustness	Robustness was determined by measuring the impact of excluding individuals who may not have understood or engaged with the survey, those who found the survey difficult, those who took a very long or short time to complete it, those who selected to treat the same patient group for all questions, and exclusion of the first and last question (susceptible to learning or fatigue effects).	"The exclusions affected the magnitude of all coefficients but only affected their significance for a small number of models containing burden of illness (BOI), where the impact was not systematic."
Ali 2017	Robustness of results to framing effects	Investigators tested four different cognitive biases based on question framing and presentation: (1) small vs. unrealistically large health inequality reductions (2) population-level vs. individual-level descriptions of health inequality reduction (3) concrete vs. abstract intervention scenarios (4) online vs. face-to-face mode of administration.	The findings of health inequality aversion are robust to framing effects. Median respondents in all scenarios displayed substantial health inequality aversion.
Li 2022	Robustness	Investigators tested whether results were robust to the inclusion of controls for gender and census region.	Results were robust.
Baker 2010	Robustness analysis of methods	Various methods were used in the relativities study to determine if the results would be robust to variations in method.	Patterns in the data were robust to the choice of method.
Comerford 2023	Robustness of perceived health inequality aversion	Controls were included in the model.	Results were robust to the inclusion of controls.
Reckers-Droog 2021	Robustness	Sensitivity analyses - excluding respondents who reported a lower clarity and certainty score for the choice task, who took an extremely long or short time to complete the task, and who stated a lower willingness-to-pay for larger quality of life gains.	"The sensitivity analyses indicated that respondents’ stated WTP in the practice task had a marginal effect on the stated WTP in the subsequent tasks (models 1 to 7: b 0.01) and that [the] results were robust."

DISCUSSION

This systematic review investigated the psychometric properties of equity-efficiency trade-off studies. Of the 115 included studies, only 33 (28.7%) assessed or discussed psychometric properties. Less than 15% of the studies assessed validity and less than 10% of the studies assessed reliability, framing and cognitive effects, or robustness. In the studies that assessed psychometric properties, investigators used a wide array of quantitative and qualitative methods to assess and report psychometric properties. To ascertain reliability, investigators commonly calculated a measure of internal consistency or offered a repeated task. Several methods were used to determine validity, including comparing results to *a priori* expectations and stated preferences to revealed preferences. Generally, equity-efficiency trade-off studies reported good reliability and validity, with fewer studies reporting low validity and reliability. Multiple studies reported no framing and cognitive effects, while others reported that their results were significantly affected by framing and cognitive effects. Most studies that assessed robustness noted that their results were robust to inclusion and exclusion of specified participants, larger sample sizes, and variations in method.

To our knowledge, this is the first systematic review to discuss the psychometric properties of equity-efficiency trade-off studies. Several review papers have assessed the validity of DCEs and other preference elicitation techniques. Merlo and colleagues conducted a validity assessment of DCEs of primary healthcare professionals, with specific assessment of internal and external validity.[137] A meta-analysis by Quaife and colleagues assessed external validity of health-related DCEs to determine how well DCEs predict health choices in real life.[13] Other reviews have investigated validity and concordance of DCEs and best-worst scaling

methods in health,[20] participant understanding in health-related DCEs,[138] and external validity of DCEs in health economics.[14] A systematic review by Ryan and colleagues described preference elicitation techniques in the health literature, including their validity, reliability, and generalizability.[139] Although these reviews discuss the psychometric properties of preference elicitation techniques, they review health-related DCEs more broadly, without a specific focus on equity as a dimension of choice.

It is unsurprising that few equity-efficiency trade-off studies report psychometric properties. A previous systematic review of self-report research utilization measures in healthcare found that only 33% of studies assessed reliability.[140] The authors of this review noted that there is significant underdevelopment in psychometric assessment.[140] The findings of our systematic review support this conclusion. It appears that psychometric properties are underreported across health research disciplines.

Previous studies have reported mixed findings with regards to the validity and reliability of preference elicitation techniques. A systematic review and meta-analysis of external validity of DCEs found that DCEs have moderate, but not exceptional, accuracy in predicting health choices.[13] Conversely, a preference elicitation study for health states that used visual analogue scale, time trade-off, and standard gamble methods found that these methods had poor test-retest reliability and construct validity.[141] Another study showed that DCEs are able to accurately predict health choices, mimicking real-world decisions, if scale and preference heterogeneity are considered.[142] Given these opposing findings in the health economics literature, it is unsurprising that we found varying reports of validity and reliability in equity-efficiency trade-off studies, with some studies reporting high or moderate validity and reliability

and other studies reporting poor psychometric properties. These conflicting findings may be due to differing preference elicitation methods, study samples, or other context-specific factors.

Implications

Only a small proportion of equity-efficiency trade-off studies assessed or discussed the psychometric properties of their measurement instruments. This finding is concerning, given that the purpose of equity-efficiency trade-off studies is to inform high-impact health policy and health resource allocation decisions. Researchers in other fields have noted that reporting psychometric properties is necessary to evaluate the usefulness of measurement instruments and meaningfully apply results to clinical practice and policy.[143] Equity-efficiency trade-off study results should only be applied to health policy if researchers can demonstrate that their results are valid and reliable. The results of these studies must be interpreted critically and cautiously if there is no assessment of psychometric properties.

Researchers should assess psychometric properties every time they are conducting a new equity-efficiency trade-off study, even if they are using a preference elicitation technique for which validity and reliability have previously been established. Psychometric properties are not fixed and depend on study context, study population, and other factors.[11] Although several studies have assessed the validity of common preference elicitation techniques such as discrete choice experiments,[13,137] these assessments are not specific to the equity-efficiency trade-off. When conducting an equity-efficiency trade-off study, it is not sufficient to cite previous studies and suggest that all discrete choice experiments are valid. The validity and reliability of

preference elicitation techniques will vary greatly depending on the context in which they are applied.

Limitations

A limitation of this systematic review is that the assessment methods of included studies were heterogeneous. Studies assessed psychometric properties using different methods and reported a wide range of qualitative and quantitative measures of psychometric properties. Moreover, studies varied widely in their preference elicitation techniques, with methods ranging from discrete choice experiments to person trade-off to integrated citizens juries and others. This heterogeneity of study methods precluded a meta-analysis or a direct comparison of psychometric properties across equity-efficiency trade-off studies. This review serves as a descriptive paper to outline the current state of psychometric property assessment and reporting in the equity-efficiency literature.

Another limitation of this systematic review is that study reviewers were not blinded to study outcomes. As such, a selection bias is possible if reviewers were more inclined to include a study if it mentioned psychometric properties in the title or abstract. Although such a selection bias is possible, it was likely mitigated due to duplicate study screening by two independent reviewers and resolution of conflicts by a third reviewer.

Suggestions for Future Research

Currently, researchers use a wide array of preference elicitation techniques in their equity-efficiency trade-off studies, such as discrete choice experiments, person trade-off, best-worst scaling, and many others. Future studies should compare the psychometric properties of

these various preference elicitation methods directly to ascertain whether certain preference elicitation methods are more valid and reliable than others in deriving preferences for equity and efficiency.

Another promising avenue for future research is determining a 'threshold' for validity and reliability in equity-efficiency trade-off studies, if such a threshold exists. Researchers and policy makers should consider how high validity and reliability of an equity-efficiency trade-off need to be for this study to be considered acceptable for use in health policy decision making.

Researchers may be unaware about the need to assess the psychometric properties of their methods. As such, future work should focus on developing guidelines to support researchers in the transparent assessment and reporting of validity and reliability in their studies. Such guidelines may outline minimum standards for the reporting of psychometric properties in preference elicitation studies.

CONCLUSION

This systematic review described psychometric properties in equity-efficiency trade-off studies. Most equity-efficiency trade-off studies did not assess reliability, validity, framing and cognitive effects, or robustness. The studies which assessed psychometric properties presented a wide range of qualitative and quantitative metrics, with most studies reporting good validity and reliability and some reporting low validity and reliability. The results of equity-efficiency trade-off studies should be interpreted with caution if there is no assessment of psychometric properties. Equity-efficiency trade-off studies should only be used to inform high-impact health policy decisions if authors demonstrate acceptable validity and reliability.

REFERENCES

1. Ali S, Ronaldson S. Ordinal preference elicitation methods in health economics and health services research: using discrete choice experiments and ranking methods. *Br Med Bull*. 2012;103(1):21-44. doi:10.1093/bmb/lds020

2. Bridges JFP. Stated preference methods in health care evaluation: an emerging methodological paradigm in health economics. *Appl Health Econ Health Policy*. 2003;2(4):213-224.

3. Louviere JJ, Flynn TN, Carson RT. Discrete Choice Experiments Are Not Conjoint Analysis. *J Choice Model*. 2010;3(3):57-72. doi:10.1016/S1755-5345(13)70014-9

4. Asamani JA, Alugsi SA, Ismaila H, Nabyonga-Orem J. Balancing Equity and Efficiency in the Allocation of Health Resources—Where Is the Middle Ground? *Healthcare*. 2021;9(10):1257. doi:10.3390/healthcare9101257

5. Culyer T. Efficiency, equity and equality in health and health care. *Effic Equity Equal Health Health Care*. Published online December 2015:1-20.

6. Cookson R, Drummond M, Weatherly H. Explicit incorporation of equity considerations into economic evaluation of public health interventions. *Health Econ Policy Law*. 2009;4(2):231-245. doi:10.1017/S1744133109004903

7. McNamara S, Holmes J, Stevely AK, Tsuchiya A. How averse are the UK general public to inequalities in health between socioeconomic groups? A systematic review. *Eur J Health Econ*. 2020;21(2):275-285. doi:10.1007/s10198-019-01126-2

8. Khor S, Elsisi ZA, Carlson JJ. How Much Does the US Public Value Equity in Health? A Systematic Review. *Value Health*. 2023;26(3):418-426. doi:10.1016/j.jval.2022.08.009

9. Cadham CJ, Prosser LA. Eliciting Trade-Offs Between Equity and Efficiency: A Methodological Scoping Review. *Value Health*. 2023;26(6):943-952. doi:10.1016/j.jval.2023.02.006

10. Asunta P, Viholainen H, Ahonen T, Rintala P. Psychometric properties of observational tools for identifying motor difficulties – a systematic review. *BMC Pediatr*. 2019;19(1):322. doi:10.1186/s12887-019-1657-6

11. Souza AC de, Alexandre NMC, Guirardello E de B. Psychometric properties in instruments evaluation of reliability and validity. *Epidemiol E Serv Saude Rev Sist Unico Saude Bras*. 2017;26(3):649-659. doi:10.5123/S1679-49742017000300022

12. Mariel P, Hoyos D, Meyerhoff J, et al. Validity and Reliability. In: Mariel P, Hoyos D, Meyerhoff J, et al., eds. *Environmental Valuation with Discrete Choice Experiments: Guidance on Design, Implementation and Data Analysis*. SpringerBriefs in Economics. Springer International Publishing; 2021:111-123. doi:10.1007/978-3-030-62669-3_8

13. Quaife M, Terris-Prestholt F, Di Tanna GL, Vickerman P. How well do discrete choice experiments predict health choices? A systematic review and meta-analysis of external validity. *Eur J Health Econ*. 2018;19(8):1053-1066. doi:10.1007/s10198-018-0954-6

14. Lancsar E, Swait J. Reconceptualising the External Validity of Discrete Choice Experiments. *PharmacoEconomics*. 2014;32(10):951-965. doi:10.1007/s40273-014-0181-7

15. Louviere JJ, Lancsar E. Choice experiments in health: the good, the bad, the ugly and toward a brighter future. *Health Econ Policy Law*. 2009;4(Pt 4):527-546. doi:10.1017/S1744133109990193

16. Dolan P, Shaw R, Tsuchiya A, Williams A. QALY maximisation and people's preferences: a methodological review of the literature. *Health Econ*. 2005;14(2):197-208. doi:10.1002/hec.924

17. Gu Y, Lancsar E, Ghijben P, Butler JR, Donaldson C. Attributes and weights in health care priority setting: A systematic review of what counts and to what extent. *Soc Sci Med*. 2015;146:41-52. doi:10.1016/j.socscimed.2015.10.005

18. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev*. 2021;10(1):89. doi:10.1186/s13643-021-01626-4

19. Ali S, Tsuchiya A, Asaria M, Cookson R. How Robust Are Value Judgments of Health Inequality Aversion? Testing for Framing and Cognitive Effects. *Med Decis Mak Int J Soc Med Decis Mak*. 2017;37(6):635-646. doi:10.1177/0272989X17700842

20. Whitty JA, Oliveira Gonçalves AS. A Systematic Review Comparing the Acceptability, Validity and Concordance of Discrete Choice Experiments and Best–Worst Scaling for Eliciting Preferences in Healthcare. *Patient - Patient-Centered Outcomes Res*. 2018;11(3):301-317. doi:10.1007/s40271-017-0288-y

21. Joy SM, Little E, Maruthur NM, Purnell TS, Bridges JFP. Patient Preferences for the Treatment of Type 2 Diabetes: A Scoping Review. *PharmacoEconomics*. 2013;31(10):877-892. doi:10.1007/s40273-013-0089-7

22. Petrou S, Kandala NB, Robinson A, Baker R. A person trade-off study to estimate age-related weights for health gains in economic evaluation. *PharmacoEconomics*. 2013;31(10):893-907. doi:10.1007/s40273-013-0085-y

23. Ahlert M, Schwettmann L. Allocating health care resources: a questionnaire experiment on the predictive success of rules. *Int J Equity Health*. 2017;16(1):112. doi:10.1186/s12939-017-0611-1

24. Whitty JA, Ratcliffe J, Chen G, Scuffham PA. Australian Public Preferences for the Funding of New Health Technologies: A Comparison of Discrete Choice and Profile Case Best-Worst Scaling Methods. *Med Decis Mak*. 2014;34(5):638-654. doi:10.1177/0272989X14526640

25. Jehu-Appiah C, Baltussen R, Acquah C, et al. Balancing equity and efficiency in health priorities in Ghana: the use of multicriteria decision analysis. *Value Health J Int Soc Pharmacoeconomics Outcomes Res*. 2008;11(7):1081-1087. doi:10.1111/j.1524-4733.2008.00392.x

26. Lancsar E, Wildman J, Donaldson C, Ryan M, Baker R. Deriving distributional weights for QALYs through discrete choice experiments. *J Health Econ*. 2011;30(2):466-478. doi:10.1016/j.jhealeco.2011.01.003

27. Ubel PA, Loewenstein G. Distributing scarce livers: the moral reasoning of the general public. *Soc Sci Med 1982*. 1996;42(7):1049-1055.

28. Cookson R, Ali S, Tsuchiya A, Asaria M. E-learning and health inequality aversion: A questionnaire experiment. *Health Econ*. 2018;27(11):1754-1771. doi:10.1002/hec.3799

29. Rowen D, Brazier J, Mukuria C, et al. Eliciting Societal Preferences for Weighting QALYs for Burden of Illness and End of Life. *Med Decis Mak Int J Soc Med Decis Mak*. 2016;36(2):210-222. doi:10.1177/0272989X15619389

30. Robson M, Asaria M, Cookson R, Tsuchiya A, Ali S. Eliciting the Level of Health Inequality Aversion in England. *Health Econ*. 2017;26(10):1328-1334. doi:10.1002/hec.3430

31. Stolk EA, Pickee SJ, Ament AHJA, Busschbach JJV. Equity in health care prioritisation: an empirical inquiry into social value. *Health Policy Amst Neth*. 2005;74(3):343-355.

32. Reckers-Droog V, van Exel J, Brouwer W. Equity Weights for Priority Setting in Healthcare: Severity, Age, or Both?. *Value Health J Int Soc Pharmacoeconomics Outcomes Res*. 2019;22(12):1441-1449. doi:10.1016/j.jval.2019.07.012

33. Li J, Casalino L, Fisman R, Kariv S, Markovits D. Experimental evidence of physician social preferences. *Proc Natl Acad Sci U S A*. 2022;119(28). doi:10.1073/pnas.2112726119

34. Ali S, Tsuchiya A, Asaria M, Cookson R. How Robust Are Value Judgments of Health Inequality Aversion? Testing for Framing and Cognitive Effects. *Med Decis Mak Int J Soc Med Decis Mak*. 2017;37(6):635-646. doi:10.1177/0272989X17700842

35. Ubel P. How stable are people's preferences for giving priority to severely ill patients? *Soc Sci Med*. 1999;49(7):895-903. doi:10.1016/S0277-9536(99)00174-4

36. Green C. Investigating public preferences on "severity of health" as a relevant condition for setting healthcare priorities. *Soc Sci Med*. 2009;68(12):2247-2255. doi:10.1016/j.socscimed.2009.03.020

37. Abásolo I, Tsuchiya A. Is more health always better for society? Exploring public preferences that violate monotonicity. *THEORY Decis*. 2013;74(4):539-563. doi:10.1007/s11238-011-9292-1

38. Werner P. Israeli lay persons' views on priority-setting criteria for Alzheimer's disease. *Health Expect Int J Public Particip Health Care Health Policy*. 2009;12(2):187-196. doi:10.1111/j.1369-7625.2008.00523.x

39. Ubel PA, Baron J, Asch DA. Preference for equity as a framing effect. *Med Decis Mak Int J Soc Med Decis Mak*. 2001;21(3):180-189.

40. Whitty J, Ratcliffe J, Kendall E, et al. Prioritising patients for bariatric surgery: building public preferences from a discrete choice experiment into public policy. *BMJ OPEN*. 2015;5(10). doi:10.1136/bmjopen-2015-008919

41. Gulácsi L, Hajdú O. Prioritization Preferences among General Practitioners in Hungary. *EAST Eur Polit Soc*. 2012;26(1):20-32. doi:10.1177/0888325410390207

42. Luyten J, Kessels R, Desmet P, Goos P, Beutels P. Priority-Setting and Personality: Effects of Dispositional Optimism on Preferences for Allocating Healthcare Resources. *Soc JUSTICE Res*. 2019;32(2):186-207. doi:10.1007/s11211-019-00329-5

43. Ubel PA, Loewenstein G. Public perceptions of the importance of prognosis in allocating transplantable livers to children. *Med Decis Mak Int J Soc Med Decis Mak*. 1996;16(3):234-241.

44. Ratcliffe J. Public preferences for the allocation of donor liver grafts for transplantation. *Health Econ*. 2000;9(2):137-148.

45. Comerford D, Tufte-Hewett A, Bridger E. Public preferences to trade-off gains in total health for health equality: Discrepancies between an abstract scenario versus the real-world scenario presented by COVID-19. *Ration Soc*. Published online 2023. doi:10.1177/10434631231193599

46. van Exel J, Baker R, Mason H, Donaldson C, Brouwer W. Public views on principles for health care priority setting: findings of a European cross-country study using Q methodology. *Soc Sci Med 1982*. 2015;126(ut9, 8303205):128-137. doi:10.1016/j.socscimed.2014.12.023

47. Schoon R, Chi C, Liu T. Quantifying public preferences for healthcare priorities in Taiwan through an integrated citizens jury and discrete choice experiment. *Soc Sci Med*. 2022;315. doi:10.1016/j.socscimed.2022.115404

48. Mckie J, Richardson J. Social preferences for prioritising the treatment of disabled and chronically ill patients: beyond the order effect. *Health Econ POLICY LAW*. 2019;14(4):443-467. doi:10.1017/S1744133118000154

49. Aidem JM. Stakeholder views on criteria and processes for priority setting in Norway: a qualitative study. *Health Policy Amst Neth*. 2017;121(6):683-690. doi:10.1016/j.healthpol.2017.04.005

50. Attema A, Brouwer W, Pinto J. The Role of Perceived Utility of Full Health in Age Weighting. *VALUE Health*. 2022;25(9):1559-1565. doi:10.1016/j.jval.2022.04.1733

51. Damschroder L, Baron J, Hershey J, Asch D, Jepson C, Ubel P. The validity of person tradeoff measurements: Randomized trial of computer elicitation versus face-to-face interview. *Med Decis Making*. 2004;24(2):170-180. doi:10.1177/0272989X04263160

52. Baltussen R, Stolk E, Chisholm D, Aikins M. Towards a multi-criteria approach for priority setting: an application to Ghana. *Health Econ*. 2006;15(7):689-696.

53. Baker R, Bateman I, Donaldson C, et al. Weighting and valuing quality-adjusted life-years using stated preference methods: preliminary results from the Social Value of a QALY Project. *Health Technol Assess*. 2010;14(27):1-+. doi:10.3310/hta14270

54. Reckers-Droog V, van Exel J, Brouwer W. Willingness to Pay for Health-Related Quality of Life Gains in Relation to Disease Severity and the Age of Patients. *Value Health J Int Soc Pharmacoeconomics Outcomes Res*. 2021;24(8):1182-1192. doi:10.1016/j.jval.2021.01.012

55. Dolan P, Cookson R. A qualitative study of the extent to which health gain matters when choosing between groups of patients. *Health POLICY*. 2000;51(1):19-30. doi:10.1016/S0168-8510(99)00079-2

56. Pinho M, Borges AP. A Scoring Index of Prioritization Factors Between Patients: An Iberian Comparison. *Health Care Manag Frederick*. 2019;38(3):267-275. doi:10.1097/HCM.0000000000000274

57. Diederich A, Winkelhage J, Wirsik N. Age as a Criterion for Setting Priorities in Health Care? A Survey of the German Public View. *PLOS ONE*. 2011;6(8). doi:10.1371/journal.pone.0023930

58. Richardson J, Mckie J, Iezzi A, Maxwell A. Age Weights for Health Services Derived from the Relative Social Willingness-to-Pay Instrument. *Med Decis Making*. 2017;37(3):239-251. doi:10.1177/0272989X16645576

59. Grepin KA, Pinkstaff CB, Hole AR, et al. Allocating external financing for health: a discrete choice experiment of stakeholder preferences. *Health Policy Plan*. 2018;33(suppl_1):i24-i30. doi:10.1093/heapol/czx017

60. Lindholm L, Rosen M, Emmelin M. An epidemiological approach towards measuring the trade-off between equity and efficiency in health policy. *Health Policy Amst Neth*. 1996;35(3):205-216.

61. Schwappach D. Are preferences for equality a matter of perspective? *Med Decis Making*. 2005;25(4):449-459. doi:10.1177/0272989X05276861

62. van de Wetering EJ, van Exel NJA, Rose JM, Hoefman RJ, Brouwer WBF. Are some QALYs more equal than others?. *Eur J Health Econ HEPAC Health Econ Prev Care*. 2016;17(2):117-127. doi:10.1007/s10198-014-0657-6

63. Li DG, Wong GX, Martin DT, et al. Attitudes on cost-effectiveness and equity: a cross-sectional study examining the viewpoints of medical professionals. *BMJ Open*. 2017;7(7):e017251. doi:10.1136/bmjopen-2017-017251

64. Quintal C. Aversion to geographic inequality and geographic variation in preferences in the context of healthcare. *Appl Health Econ Health Policy*. 2009;7(2):121-136. doi:10.2165/10899530-000000000-00000

65. Berhane Defaye F, Danis M, Wakim P, Berhane Y, Norheim OF, Miljeteig I. Bedside Rationing Under Resource Constraints-A National Survey of Ethiopian Physicians' Use of Criteria for Priority Setting. *AJOB Empir Bioeth*. 2019;10(2):125-135. doi:10.1080/23294515.2019.1583691

66. Pinho M, Borges A, Zahariev B. Bedsides healthcare rationing dilemmas: A survey from Bulgaria and comparison with Portugal. *Soc THEORY Health*. 2017;15(3):285-301. doi:10.1057/s41285-017-0029-2

67. Diederich A, Swait J, Wirsik N. Citizen Participation in Patient Prioritization Policy Decisions: An Empirical and Experimental Study on Patients' Characteristics. *PLOS ONE*. 2012;7(5):e36824. doi:10.1371/journal.pone.0036824

68. Howard K, Jan S, Rose J, et al. Community Preferences for the Allocation &Donation of Organs - The PAraDOx Study. *BMC PUBLIC Health*. 2011;11. doi:10.1186/1471-2458-11-386

69. Browning C, Thomas S. Community values and preferences in transplantation organ allocation decisions. *Soc Sci Med*. 2001;52(6):853-861. doi:10.1016/S0277-9536(00)00187-8

70. Chim L, Salkeld G, Kelly PJ, Lipworth W, Hughes DA, Stockler MR. Community views on factors affecting medicines resource allocation: cross-sectional survey of 3080 adults in Australia. *Aust Health Rev Publ Aust Hosp Assoc*. 2019;43(3):254-260. doi:10.1071/AH16209

71. Baji P, Garcia-Goni M, Gulacsi L, Mentzakis E, Paolucci F. Comparative analysis of decision maker preferences for equity/efficiency attributes in reimbursement decisions in three European countries. *Eur J Health Econ HEPAC Health Econ Prev Care*. 2016;17(7):791-799. doi:10.1007/s10198-015-0721-x

72. Goto R, Mori T. Comparison of Equity Preferences for Life Expectancy Gains: A Discrete Choice Experiment Among the Japanese and Korean General Public. *Value Health Reg Issues*. 2019;18(101592642):8-13. doi:10.1016/j.vhri.2018.05.004

73. Youngkong S, Baltussen R, Tantivess S, Koolman X, Teerawattananon Y. Criteria for priority setting of HIV/AIDS interventions in Thailand: a discrete choice experiment. *BMC Health Serv Res*. 2010;10. doi:10.1186/1472-6963-10-197

74. Mirelman A, Mentzakis E, Kinter E, et al. Decision-making criteria among national policymakers in five countries: a discrete choice experiment eliciting relative preferences for equity and efficiency. *Value Health J Int Soc Pharmacoeconomics Outcomes Res*. 2012;15(3):534-539. doi:10.1016/j.jval.2012.04.001

75. Ratcliffe J, Buxton M, Young T, Longworth L. Determining priority for liver transplantation: a comparison of cost per QALY and discrete choice experiment-generated public preferences. *Appl Health Econ Health Policy*. 2005;4(4):249-255.

76. Tsuchiya A, Dolan P. Do NHS clinicians and members of the public share the same views about reducing inequalities in health?. *Soc Sci Med 1982*. 2007;64(12):2499-2503.

77. Shmueli A. Do the equity-efficiency preferences of the Israeli Basket Committee match those of Israeli health policy makers?. *Isr J Health Policy Res*. 2017;6(101584158):20. doi:10.1186/s13584-017-0145-4

78. Kolasa K, Lewandowski T. Does it matter whose opinion we seek regarding the allocation of healthcare resources? - a case study. *BMC Health Serv Res*. 2015;15(101088677):564. doi:10.1186/s12913-015-1210-8

79. Denburg A, Ungar W, Chen S, Hurley J, Abelson J. Does moral reasoning influence public values for health care priority setting?: A population-based randomized stated preference survey. *Health POLICY*. 2020;124(6):647-658. doi:10.1016/j.healthpol.2020.04.007

80. McNamara S, Tsuchiya A, Holmes J. Does the UK-public's aversion to inequalities in health differ by group-labelling and health-gain type? A choice-experiment. *Soc Sci Med 1982*. 2021;269(ut9, 8303205):113573. doi:10.1016/j.socscimed.2020.113573

81. Shmueli A, Golan O, Paolucci F, Mentzakis E. Efficiency and equity considerations in the preferences of health policy-makers in Israel. *Isr J Health Policy Res*. 2017;6(101584158):18. doi:10.1186/s13584-017-0142-7

82. Norman R, Hall J, Street D, Viney R. Efficiency and equity: a stated preference approach. *Health Econ*. 2013;22(5):568-581. doi:10.1002/hec.2827

83. Li M, Colby H, Fernbach P. Efficiency for Lives, Equality for Everything Else: How Allocation Preference Shifts Across Domains. *Soc Psychol Personal Sci*. 2019;10(5):697-707. doi:10.1177/1948550618783709

84. Lim MK, Bae EY, Choi SE, Lee EK, Lee TJ. Eliciting public preference for health-care resource allocation in South Korea. *Value Health J Int Soc Pharmacoeconomics Outcomes Res*. 2012;15(1 Suppl):S91-4. doi:10.1016/j.jval.2011.11.014

85. Gibbs N, Powell PA, Tsuchiya A. Equal access for equal need: Eliciting public preferences for access to health treatment by employment status. *Soc Sci Med 1982*. 2019;222(ut9, 8303205):246-255. doi:10.1016/j.socscimed.2019.01.003

86. Paolucci F, Mentzakis E, Defechereux T, Niessen LW. Equity and efficiency preferences of health policy makers in China--a stated preference analysis. *Health Policy Plan*. 2015;30(8):1059-1066. doi:10.1093/heapol/czu123

87. Mentzakis E., Garcia-Goni M., Sequeira A.R., Paolucci F. Equity and efficiency priorities within the Spanish health system: A discrete choice experiment eliciting stakeholders preferences. *Health Policy Technol*. 2019;8(1):30-41. doi:10.1016/j.hlpt.2019.01.003

88. Lal A, Mohebi M, Sweeney R, Moodie M, Peeters A, Carter R. Equity Weights for Socioeconomic Position: Two Methods-Survey of Stated Preferences and Epidemiological Data. *Value Health J Int Soc Pharmacoeconomics Outcomes Res*. 2019;22(2):247-253. doi:10.1016/j.jval.2018.07.006

89. van den Broek-Altenburg EM, Benson JS, Atherly AJ. Examining alignment of community health teams' preferences for health, equity, and spending with state all-payer waiver priorities: A discrete choice experiment. *Health Serv Res*. 2023;(g2l, 0053006). doi:10.1111/1475-6773.14257

90. Ratcliffe J, Bekker HL, Dolan P, Edlin R. Examining the attitudes and preferences of health care decision-makers in relation to access, equity and cost-effectiveness: a discrete choice experiment. *Health Policy Amst Neth*. 2009;90(1):45-57. doi:10.1016/j.healthpol.2008.09.001

91. McHugh N, Pinto-Prades JL, Baker R, Mason H, Donaldson C. Exploring the relative value of end of life QALYs: Are the comparators important?. *Soc Sci Med 1982*. 2020;245(ut9, 8303205):112660. doi:10.1016/j.socscimed.2019.112660

92. Green C, Gerard K. EXPLORING THE SOCIAL VALUE OF HEALTH-CARE INTERVENTIONS: A STATED PREFERENCE DISCRETE CHOICE EXPERIMENT. *Health Econ*. 2009;18(8):951-976. doi:10.1002/hec.1414

93. Lindholm L.A., Emmelin M.A., Rosen M.E. Health maximization rejected: The view of Swedish politicians. *Eur J Public Health*. 1997;7(4):405-410. doi:10.1093/eurpub/7.4.405

94. Dolan P, Tsuchiya A. Health priorities and public preferences: the relative importance of past health experience and future health prospects. *J Health Econ*. 2005;24(4):703-714. doi:10.1016/j.jhealeco.2004.11.007

95. Franken M, Koolman X. Health system goals: a discrete choice experiment to obtain societal valuations. *Health Policy Amst Neth*. 2013;112(1-2):28-34. doi:10.1016/j.healthpol.2012.12.013

96. Lindholm L, Rosen M, Emmelin M. How many lives is equity worth? A proposal for equity adjusted years of life saved. *J Epidemiol Community Health*. 1998;52(12):808-811.

97. Steinert JI, Sternberg H, Veltri GA, Buthe T. How should COVID-19 vaccines be distributed between the Global North and South: a discrete choice experiment in six European countries. *eLife*. 2022;11(101579614). doi:10.7554/eLife.79819

98. Hadian M, Jouyani Y, Asadi H, Palangi H, Rahimnia R. Identifying the Criteria Affecting Appropriate Allocation of Health System Resources to Different Diseases in Iran: A Qualitative Inquiry. *Health SCOPE*. 2019;8(4). doi:10.5812/jhealthscope.83968

99. Yang F., Katumba K.R., Greco G., et al. INCORPORATING CONCERN FOR HEALTH EQUITY INTO RESOURCE ALLOCATION DECISIONS: DEVELOPMENT OF A TOOL AND POPULATIONBASED VALUATION FOR UGANDA. *BMJ Glob Health*. 2022;7(Supplement 2):A13. doi:10.1136/bmjgh-2022-ISPH.36

100. Pinho M, Botelho A. Inference Procedures to Quantify the Efficiency-Equality Trade-Off in Health from Stated Preferences: A Case Study in Portugal. *Appl Health Econ Health Policy*. 2018;16(4):503-513. doi:10.1007/s40258-018-0394-6

101. Gyrd-Hansen D. Investigating the social value of health changes. *J Health Econ*. 2004;23(6):1101-1116.

102. Dolan P, Tsuchiya A. It is the lifetime that matters: public preferences over maximising health and reducing inequalities in health. *J Med Ethics*. 2012;38(9):571-573. doi:10.1136/medethics-2011-100228

103. Ubel P, Richardson J, Prades J. Life-saving treatments and disabilities - Are all QALYs created equal? *Int J Technol Assess Health CARE*. 1999;15(4):738-748. doi:10.1017/S0266462399154138

104. Ottersen T, Maestad O, Norheim OF. Lifetime QALY prioritarianism in priority setting: quantification of the inherent trade-off. *Cost Eff Resour Alloc CE*. 2014;12(1):2. doi:10.1186/1478-7547-12-2

105. Youngkong S, Teerawattananon Y, Tantivess S, Baltussen R. Multi-criteria decision analysis for setting priorities on HIV/AIDS interventions in Thailand. *Health Res Policy Syst*. 2012;10(101170481):6. doi:10.1186/1478-4505-10-6

106. Perneger T, Martin D, Bovier P. Physicians' attitudes toward health care rationing. *Med Decis Making*. 2002;22(1):65-70. doi:10.1177/02729890222062928

107. Jelsma J, Shumba D, Kristian H, De Weerdt W, De Cock P. Preferences of urban Zimbabweans for health and life lived at different ages. *Bull World Health Organ*. 2002;80(3):204-209.

108. Erdem S, Thompson C. Prioritising health service innovation investments using public preferences: a discrete choice experiment. *BMC Health Serv Res*. 2014;14. doi:10.1186/1472-6963-14-360

109. Olsen JA. Priority preferences: "end of life" does not matter, but total life does. *Value Health J Int Soc Pharmacoeconomics Outcomes Res*. 2013;16(6):1063-1066. doi:10.1016/j.jval.2013.06.002

110. Otim ME, Kelaher M, Anderson IP, Doran CM. Priority setting in Indigenous health: assessing priority setting process and criteria that should guide the health system to improve Indigenous Australian health. *Int J Equity Health*. 2014;13(101147692):45. doi:10.1186/1475-9276-13-45

111. Meusel V, Mentzakis E, Baji P, Fiorentini G, Paolucci F. Priority setting in the German healthcare system: results from a discrete choice experiment. *Int J Health Econ Manag*. 2023;23(3):411-431. doi:10.1007/s10754-023-09347-y

112. Wouters S, van Exel J, Baker R, B F Brouwer W. Priority to End of Life Treatments? Views of the Public in the Netherlands. *Value Health J Int Soc Pharmacoeconomics Outcomes Res*. 2017;20(1):107-117. doi:10.1016/j.jval.2016.09.544

113. Lee JS, Kim S, Do YK. Public Preferences for Allocation Principles for Scarce Medical Resources in the COVID-19 Pandemic in Korea: Comparisons With Ethicists' Recommendations. *J Prev Med Public Health Yebang Uihakhoe Chi*. 2021;54(5):360-369. doi:10.3961/jpmph.21.333

114. Luyten J, Kessels R, Goos P, Beutels P. Public preferences for prioritizing preventive and curative health care interventions: a discrete choice experiment. *Value Health J Int Soc Pharmacoeconomics Outcomes Res*. 2015;18(2):224-233. doi:10.1016/j.jval.2014.12.007

115. Bae EY, Lim MK, Lee B, Bae G, Hong J. Public preferences in healthcare resource allocation: A discrete choice experiment in South Korea. *Health Policy Amst Neth*. 2023;138(8409431, hep):104932. doi:10.1016/j.healthpol.2023.104932

116. Luyten J, Dorgali V, Hens N, Beutels P. Public preferences over efficiency, equity and autonomy in vaccination policy: An empirical study. *Soc Sci Med*. 2013;77:84-89. doi:10.1016/j.socscimed.2012.11.009

117. Darvishi A, Daroudi R, Yaseri M, Sari AA. Public preferences regarding the priority setting criteria of health interventions for budget allocation: results of a survey of Iranian adults. *BMC Public Health*. 2022;22(1):2038. doi:10.1186/s12889-022-14404-1

118. Blacksher E, Rigby E, Espey C. Public values, health inequality, and alternative notions of a "fair" response. *J Health Polit Policy Law*. 2010;35(6):889-920. doi:10.1215/03616878-2010-033

119. Luyten J, Beutels P, Vandermeulen C, Kessels R. Social preferences for adopting new vaccines in the national immunization program: A discrete choice experiment. *Soc Sci Med 1982*. 2022;303(ut9, 8303205):114991. doi:10.1016/j.socscimed.2022.114991

120. McKie J, Richardson J. Social preferences for prioritizing the treatment of severely ill patients: The relevance of severity, expected benefit, past health and lifetime health. *Health Policy Amst Neth*. 2017;121(8):913-922. doi:10.1016/j.healthpol.2017.05.010

121. Chim L, Salkeld G, Kelly P, Lipworth W, Hughes DA, Stockler MR. Societal perspective on access to publicly subsidised medicines: A cross sectional survey of 3080 adults in Australia. *PloS One*. 2017;12(3):e0172971. doi:10.1371/journal.pone.0172971

122. Skedgel C, Wailoo A, Akehurst R. Societal preferences for distributive justice in the allocation of health care resources: a latent class discrete choice experiment. *Med Decis Mak Int J Soc Med Decis Mak*. 2015;35(1):94-105. doi:10.1177/0272989X14547915

123. Linley WG, Hughes DA. Societal views on NICE, cancer drugs fund and value-based pricing criteria for prioritising medicines: a cross-sectional survey of 4118 adults in Great Britain. *Health Econ*. 2013;22(8):948-964. doi:10.1002/hec.2872

124. Arroyos-Calvera D, Covey J, Loomes G, McDonald R. The efficiency-equity trade-off, self-interest, and moral principles in health and safety valuation. *Soc Sci Med 1982*. 2019;238(ut9, 8303205):112477. doi:10.1016/j.socscimed.2019.112477

125. Skedgel C. The prioritization preferences of pan-Canadian Oncology Drug Review members and the Canadian public: a stated-preferences comparison. *Curr Oncol Tor Ont*. 2016;23(5):322-328.

126. Bae E.Y., Lim M.K., Choi S.E., Lee T.J. The public's preference on the priorities in health care. *Value Health*. 2010;13(7):A534.

127. Mason H, Baker R, Donaldson C. Understanding public preferences for prioritizing health care interventions in England: does the type of health gain matter? *J Health Serv Res POLICY*. 2011;16(2):81-89. doi:10.1258/jhsrp.2010.010039

128. Ratcliffe J, Lancsar E, Walker R, Gu Y. Understanding what matters: An exploratory study to investigate the views of the general public for priority setting criteria in health care. *Health Policy Amst Neth*. 2017;121(6):653-662. doi:10.1016/j.healthpol.2017.03.003

129. Nicolet A, van Asselt A, Vermeulen K, Krabbe P. Value judgment of new medical treatments: Societal and patient perspectives to inform priority setting in The Netherlands. *PLOS ONE*. 2020;15(7). doi:10.1371/journal.pone.0235666

130. Shah KK, Tsuchiya A, Wailoo AJ. Valuing health at the end of life: a stated preference discrete choice experiment. *Soc Sci Med 1982*. 2015;124(ut9, 8303205):48-56. doi:10.1016/j.socscimed.2014.11.022

131. van de Wetering L, van Exel J, Bobinac A, Brouwer WBF. Valuing QALYs in Relation to Equity Considerations Using a Discrete Choice Experiment. *PharmacoEconomics*. 2015;33(12):1289-1300. doi:10.1007/s40273-015-0311-x

132. Lal A, Siahpush M, Moodie M, Peeters A, Carter R. Weighting Health Outcomes by Socioeconomic Position Using Stated Preferences. *PharmacoEconomics - Open*. 2018;2(1):43-51. doi:10.1007/s41669-017-0036-1

133. Card K, Adshade M, Hogg R, Jollimore J, Lachowsky N. What public health interventions do people in Canada prefer to fund? A discrete choice experiment. *BMC PUBLIC Health*. 2022;22(1). doi:10.1186/s12889-022-13539-5

134. Jain V, Atun R, Hansen P, Lorgelly P. Which countries need COVID-19 vaccines the most? Development of a prioritisation tool. *BMC Public Health*. 2022;22(1):1518. doi:10.1186/s12889-022-13948-6

135. Bae EY, Lim MK, Lee B, Bae G. Who should be given priority for public funding?. *Health Policy Amst Neth*. 2020;124(10):1108-1114. doi:10.1016/j.healthpol.2020.06.010

136. Reckers-Droog V, van Exel J, Brouwer W. Who should receive treatment? An empirical enquiry into the relationship between societal views and preferences concerning healthcare priority setting. *PloS One*. 2018;13(6):e0198761. doi:10.1371/journal.pone.0198761

137. Merlo G, van Driel M, Hall L. Systematic review and validity assessment of methods used in discrete choice experiments of primary healthcare professionals. *Health Econ Rev*. 2020;10(1):39. doi:10.1186/s13561-020-00295-8

138. Pearce A, Harrison M, Watson V, et al. Respondent Understanding in Discrete Choice Experiments: A Scoping Review. *The Patient*. 2021;14(1):17-53. doi:10.1007/s40271-020-00467-y

139. Ryan M, Scott DA, Reeves C, et al. Eliciting public preferences for healthcare: a systematic review of techniques. *Health Technol Assess Winch Engl*. 2001;5(5):1-186. doi:10.3310/hta5050

140. Squires JE, Estabrooks CA, O'Rourke HM, Gustavsson P, Newburn-Cook CV, Wallin L. A systematic review of the psychometric properties of self-report research utilization measures used in healthcare. *Implement Sci*. 2011;6(1):83. doi:10.1186/1748-5908-6-83

141. Crump RT, Lau R, Cox E, Currie G, Panepinto J. Testing the feasibility of eliciting preferences for health states from adolescents using direct methods. *BMC Pediatr*. 2018;18(1):199. doi:10.1186/s12887-018-1179-7

142. Bekker-Grob EW de, Swait JD, Kassahun HT, et al. Are Healthcare Choices Predictable? The Impact of Discrete Choice Experiment Designs and Models. *Value Health*. 2019;22(9):1050-1062. doi:10.1016/j.jval.2019.04.1924

143. Zangaro GA. Importance of Reporting Psychometric Properties of Instruments Used in Nursing Research. *West J Nurs Res*. 2019;41(11):1548-1550. doi:10.1177/0193945919866827

Appendix A: Full Database Search Strategies

MEDLINE

1 health.tw,kf.
2 (healthcare or health-care or 'health care').tw,kf.
3 *Health Priorities/
4 *Health Care Rationing/
5 *"Value of Life"/
6 *Quality-Adjusted Life Years/
7 (QALY or "quality-adjusted life year$").tw,kf.
8 *Resource Allocation/
9 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8
10 ('choice behaviour' or 'choice behavior' or 'choice experiment*' or 'value elicit*' or DCE or 'person trade-off$' or 'person tradeoff$' or PTO or 'conjoint analysis' or 'preference* elicit*' or 'trade-off$' or 'tradeoff$').tw,kf.
11 ('stated preference*' or 'public preference*' or 'community preference*' or 'societal preference*' or 'priority setting').tw,kf.
12 ('social value$' or 'societal value$' or (distribut* adj2 preference*) or 'social choice' or 'relative value$' or 'community value$').tw,kf.
13 *Choice Behavior/
14 10 or 11 or 12 or 13
15 (inequalit* or equit* or inequit*).tw,kf.
16 ((distribution* adj weight*) or "equity weight*" or (QALY adj2 weight*) or (equity adj2 preference*) or "lifetime health" or (QALY and "relative value")).tw,kf.
17 ((health adj maximi*) or "health benefit maximi*").tw,kf.
18 ("outcome egalitaria*" or "gain egalitaria*" or prioritaria* or sufficientaria* or Rawlsian).tw,kf.
19 ('Social Welfare Function*' or SWF).tw,kf.
20 'inequality aversion'.tw,kf.
21 ("fair innings" or "egalitarian ageism" or "age-related weights" or "age-weighting preferences").tw,kf.
22 ('absolute shortfall' or 'proportional shortfall').tw,kf.
23 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22
24 9 and 14 and 23

Embase

1 health.tw,kf.
2 (healthcare or health-care or 'health care').tw,kf.
3 "health priorities".tw,kf.
4 "health care rationing".tw,kf.
5 "value of life".tw,kf.
6 "quality-adjusted life years".tw,kf. or *quality adjusted life year/
7 "resource allocation".tw,kf. or *resource allocation/
8 1 or 2 or 3 or 4 or 5 or 6 or 7
9 ('choice behaviour' or 'choice behavior' or 'choice experiment*' or 'value elicit*' or DCE or 'person trade-off$' or 'person tradeoff$' or PTO or 'conjoint analysis' or 'preference* elicit*' or 'trade-off$' or 'tradeoff$').tw,kf.
10 ('stated preference*' or 'public preference*' or 'community preference*' or 'societal preference*' or 'priority setting').tw,kf.
11 ('social value$' or 'societal value$' or (distribut* adj2 preference*) or 'social choice' or 'relative value$' or 'community value$').tw,kf.
12 9 or 10 or 11
13 (inequalit* or equit* or inequit*).tw,kf.
14 ((distribution* adj weight*) or "equity weight*" or (QALY adj2 weight*) or (equity adj2 preference*) or "lifetime health" or (QALY and "relative value")).tw,kf.
15 ((health adj maximi*) or "health benefit maximi*").tw,kf.
16 ("outcome egalitaria*" or "gain egalitaria*" or prioritaria* or sufficientaria* or Rawlsian).tw,kf.

17 ('Social Welfare Function*' or SWF).tw,kf.
18 'inequality aversion'.tw,kf.
19 ("fair innings" or "egalitarian ageism" or "age-related weights" or "age-weighting preferences").tw,kf.
20 ('absolute shortfall' or 'proportional shortfall').tw,kf.
21 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20
22 8 and 12 and 21

Web of Science

"health priorit*" or "health$care rationing" or "value of life" or "quality-adjusted life year$" or "resource allocation" or "health" or "health$care" or "health benefit*"

AND

"choice behavior" or "choice behaviour" or "choice experiment*" or "value elicit*" or DCE or "person trade-off$" or "person tradeoff$" or PTO or "conjoint analysis" or "preference* elicit*" or "trade-off$" or "tradeoff$" or "stated preference*" or "public preference*" or "community preference*" or "societal preference*" or "priority setting" or "social value$" or "societal value$" or (distribut* adj2 preference*) or "social choice" or "relative value$" or "community value$" or "rationing scenario" or "attitude to health"

AND

"equalit*" or "inequalit*" or "equit*" or "inequit*" or (distribution* adj weight*) or "equity weight*" or (QALY adj2 weight*) or (equity adj2 preference*) or (QALY and relative value) or (health adj maximi*) or "health benefit maximi*" or egalitaria* or prioritaria* or sufficientaria* or "health inequality aversion" or "social welfare function" or SWF or "fair innings" or "ageism" or "age-related weights" or "age-weighting preferences" or "absolute shortfall" or "proportional shortfall" or (preference adj3 "severely illl") or "health prospect*" or "health status disparities" or "severity of illness index" or "socioeconomic factors"