



REVIEW ARTICLE



A systematic review of direct preference measurements in health states treated with plastic surgery

Emma Hansson^{a,b} , Lars Sandman^{c,d,e}  and Thomas Davidson^c 

^aDepartment of Plastic and Reconstructive Surgery, Sahlgrenska University Hospital, Gothenburg, Sweden; ^bDepartment of Clinical Sciences, The Sahlgrenska Academy, Gothenburg University, Gothenburg, Sweden; ^cNational Centre for Priorities in Health, Department of Medical and Health Sciences, Linköping University, Linköping, Sweden; ^dVästra Götaland Region, Gothenburg, Sweden; ^eFaculty of Police Work, Department of Campus Police Education, Borås University, Borås, Sweden

ABSTRACT

One way to compare health care needs and outcomes on common scales is by estimating the strength of preferences or willingness-to-pay (WTP). The aim of this study was to review directly measured preference values and WTP estimates for health states treated by plastic surgery. The included articles had to meet the criteria defined in the SPIDER (Sample, Phenomenon of Interest, Design, Evaluation, Research type). Relevant databases were searched using predetermined strings. Data were extracted in a standardised manner. Included studies were appraised according to the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach for rating the importance of outcomes. In total, 213 abstracts were retrieved. Of these, 179 did not meet the inclusion criteria and were excluded, leaving 34 studies in the review. The risk of bias was considered moderate in four studies and serious in the rest. The overall certainty of evidence for directly measured preference values and WTP estimates for health states treated by plastic surgery is low (Grade $\Theta\Theta\Theta\Theta$). The lowest preference scores were generally elicited for facial defects/anomalies and the highest for excess skin after massive weight loss. Scientific knowledge about preferences and the resulting health gains might play an essential role in deciding which procedures should be considered for public funding or rather rationed within the system. Better quality studies are required to allow for such applications.

Abbreviations: DCE: discrete choice experiment; DIEP: deep inferior epigastric perforator flap; EQ: 5D: EuroQoL gen pop general public; GRADE: grading of recommendations assessment, development, and evaluation; Med exp: medical expert; Med stud: medical student; MSW: massive weight loss; Pat: patients with the condition in question; QALY: quality-adjusted life years; RCT: randomised controlled trial; SF: short form health survey; SG: standard gamble; SPIDER: sample, phenomenon of interest, design, evaluation, research type; TRAM: transverse rectus abdominis myocutaneous flap; TTO: time trade-off; VAS: visual analogue scale; WTP: Willingness-to-pay

ARTICLE HISTORY

Received 7 December 2020
Revised 28 May 2021
Accepted 1 July 2021

KEYWORDS



Plastic surgery; rationing; prioritizing; health care need; preferences; contingent valuation; willingness-to-pay


Introduction

Health care systems are increasingly struggling with balancing limited resources, given demographic and technological developments, and increased expectations. The nature of plastic surgery entails an element of subjectivity when it comes to deciding who should receive treatment, how conditions should be treated, and whether the outcome is successful. In fact, studies have revealed that there is a variation in the conditions that are treated with publicly funded plastic surgery, indicating a lack of evidence and unequal distribution of health care resources [1–3]. Therefore, a more consistent approach to which plastic surgical procedures should be granted, and to whom, within publicly funded systems is needed.

One way to compare health care needs and outcomes is to estimate the strength of preferences, a measure of how people

value different health states on a common interval scale that is often between 0 and 1, where 0 indicates death and 1 indicates perfect health. A preferred health state will receive a greater value. Preference values can be measured from different perspectives, such as that of the public ('societal score'), patients (valuing their own or hypothetical medical conditions), and medical experts. Preference values can be elicited directly with measurements, such as with the Visual Analogue Scale (VAS), Time Trade-Off (TTO), Standard Gamble (SG), and Discrete Choice Experiment (DCE), or indirectly with generic or disease-specific quality of life instruments, such as the EuroQoL (EQ-5D), Short Form Health Survey (SF-36 or SF-6D [dimensions]), Health Utility Index, and Breast-Q [4]. Indirect measures are often used, as it can be time consuming, difficult, and sometimes unethical to measure preferences using direct methods. This is often the case in large randomised controlled trials (RCTs) comparing different treatment

CONTACT Emma Hansson  emma.hansson.2@gu.se  Department of Plastic and Reconstructive Surgery, Sahlgrenska University Hospital, Gröna Stråket 8, Gothenburg, SE-413 45, Sweden
The work has never been presented previously.

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/2000656X.2021.1953039>

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

methods, where preferences are calculated on multiple occasions in many patients [5]. However, when preferences are used to study healthcare needs in a priority setting, direct measures are preferable.

Preference measures can be combined with survival estimates to calculate quality-adjusted life years (QALY), a measure in which both the effects of health care intervention on the patient's quality of life and mortality are combined. QALYs are calculated by multiplying the time spent in a health state by the preference value for that health state. QALYs can be seen as a 'common currency' that enables comparison between different health states and different areas of health care [4,6,7]. However, when healthcare resources are allocated, people's preferences for a certain health state are taken into consideration, as well as people's preferences for prioritising health care for the severely ill and the socioeconomically disadvantaged [7]. In a needs-based healthcare system, it is generally the case that a greater need results in a greater claim to public funding (*ceteris paribus*). Recent conceptual developments concerning healthcare need define it as a combination of the severity of the condition, that is, how bad a patient is and the benefit an intervention can bring the patient [8].

Another method to compare different health states on a common scale is to estimate the willingness-to-pay (WTP), which represents the amount of resources the respondent (e.g. member of society/patient) is maximally willing to allocate to a certain treatment, that is, to acquire or avoid a certain health state. WTP is often established with a bidding method, that is, respondents are asked if they would be willing to pay more than a stated value (yes/no) for a certain procedure. The process is iterated for a set number of times to obtain the final WTP value.

The aims of this study were to review directly measured preference values and WTP estimates for health states treated by plastic surgery and to establish the risk of bias of the studies.

Methods

Protocol

The study protocol was registered in Prospero (CRD42020154101).

Eligibility criteria

The inclusion criteria were studies examining preference-based measures and WTP in plastic surgery, as measured with direct methodology. The included articles had to meet the criteria defined in a SPIDER (Sample, Phenomenon of Interest, Design, Evaluation, Research type) [9]. Sample: Health states treated with plastic surgery. All domains of plastic surgery, as defined by the European Training Requirements for the specialty of Plastic, Reconstructive, and Aesthetic Surgery of the European Union of Medical Specialists, except hand surgery, were included [10]. Hand surgery was excluded because it is only a part of plastic surgery in some countries. Phenomenon of interest: Preferences, which is a measure of how people value different health states. Design: Preferences measured with TTO, SG, VAS, DCE, or WTP determined using methods such as the contingent valuation method and/or the iterative bidding method. Evaluation: Preferences elicited from the general public, patients, and other stakeholders Research type: RCTs, non-RCTs, observational studies, economical evaluations, and case series were included. Two of the authors (EH and TD) independently assessed whether the articles met the inclusion criteria, and disagreements were resolved by discussion.

Information sources, search, and study selection

Medline, PubMed, CINAHL, Embase, Cochrane library, and EconLit databases were searched for articles and abstracts published between January 1950 and June 2020. The search was performed on 20/07/15. No grey literature sources were identified. The search strings were ((((((plastic surgery) OR (aesthetic surgery)) OR (aesthetic surgery)) OR (cosmetic surgery)) OR (reconstructive surgery)) AND (((((time trade-off) OR (standard gamble) OR (discrete choice experiment*)) OR (willingness to pay))) AND (((((QALY) OR (QALY weight*)) OR (HRQoL weight*)) OR (utility outcome score*)) OR (utility score*)) OR (preference-based measurement*)) OR (contingent valuation)) in Medline/PubMed. In the other databases, the search string ((((((plastic surgery) OR (aesthetic surgery)) OR (aesthetic surgery)) OR (cosmetic surgery)) OR (reconstructive surgery)) AND (((((time trade-off) OR (standard gamble) OR (discrete choice experiment*)) OR (willingness to pay))) was used. Moreover, all bibliographies of the included studies were checked manually. The search was limited to studies published in English, French, German, Italian, Swedish, Danish, and Norwegian. When eligibility for inclusion could not be assessed by reading the abstract, the entire article was read and assessed.

Data collection process and data items

Information collected included the first author, year of publication, study country, perspective, and sample size, excluding respondents, health states, preference measurements, WTP measurements, number of clinical scenarios, if the description of the health states and scenarios were presented, if there were examples of photos illustrating scenarios included in the article, how the measurements were performed (computer-based/interviews), if control scenarios with healthy subjects and blindness were included, if the surveys were field tested before use, and if the demography of the responders was presented.

Risk of bias in individual studies and across studies and certainty of evidence

We used the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach to assess the certainty of evidence for rating the importance of outcomes [11,12]. The overall risk of bias of individual studies was assessed across the subdomains developed by the GRADE working group as 'low', 'moderate', 'serious', and 'critical'. The evaluated domains included selection of participants into the study, completeness of data, measurement instrument (choice of instrument, administration of instrument, outcome presentation, and understanding of the instrument by the study population), data analysis, and publication bias (bias arising from expertise and funding sources). The risk of bias for the body of evidence was assessed as 'not serious', 'serious', or 'very serious' [11].

The overall certainty of evidence was rated down based on the assessment of risk of bias, indirectness (due to SPIDER elements and due to methodological elements) [11], inconsistency, and imprecision [12] and finally rated as 'High' (⊕⊕⊕⊕), 'Moderate' (⊕⊕⊕⊕), 'Low' (⊕⊕⊕⊕), or 'Very low' (⊕⊕⊕⊕) [13].

Results

Study selection

A total of 213 abstracts were retrieved following the search and manual checks of the bibliographies (Figure 1). Of these, 145 did not

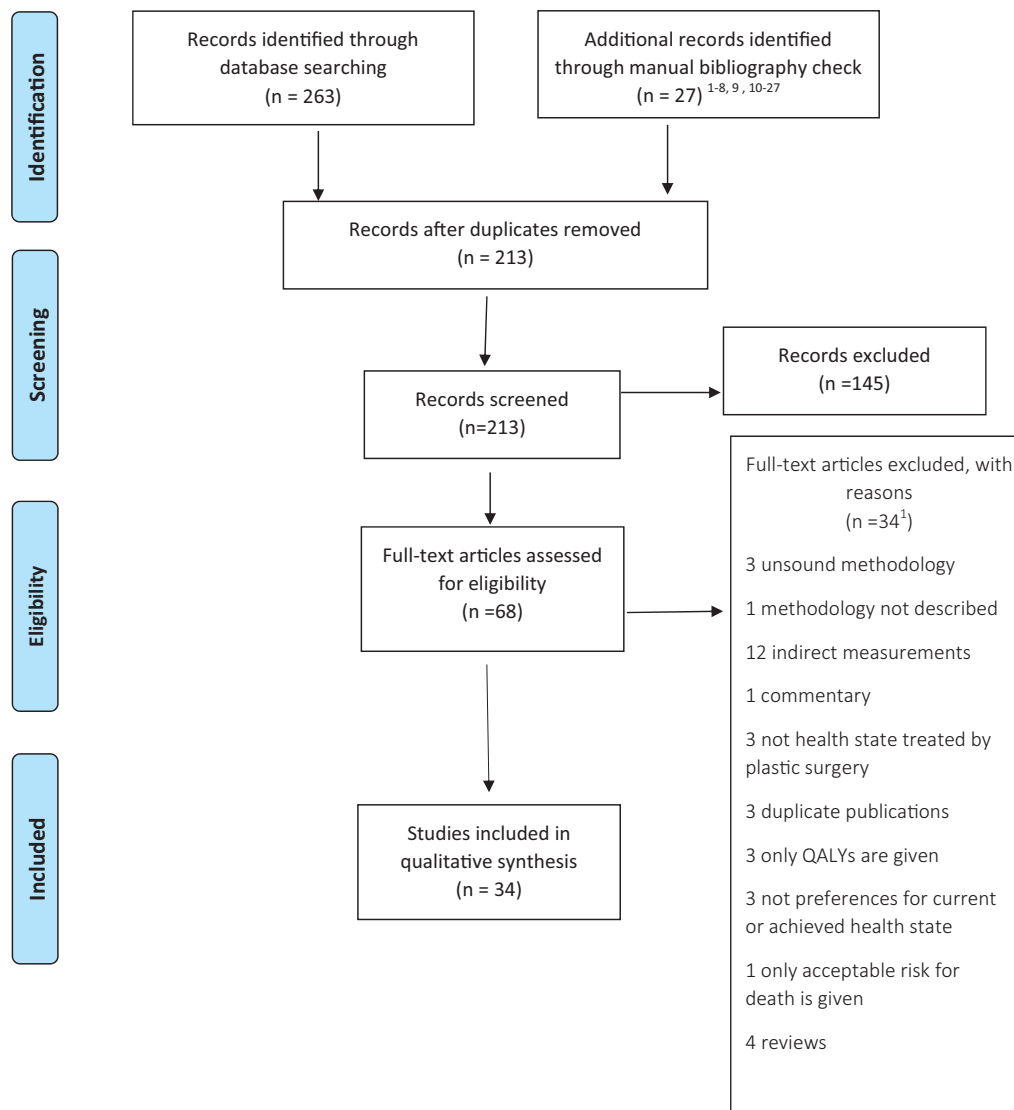


Figure 1. PRISMA flow diagram.

meet the inclusion criteria and were excluded, leaving 68 articles that were read in full text. After a more detailed examination, 34 articles were excluded (Supplement 1), leaving 34 studies to be included in the review (Table 1). Twenty-two studies used three direct methods (VAS, TTO, and SG) to elicit preference values, and ten studies used one or two of the methods (Table 1). None of the studies used DCE. Five studies examined the WTP (Table 1).

Risk of bias within and across studies

Regarding the selection of participants into the study, preferences were identified from the general public in twenty-eight studies, from students in eight studies,¹ from patients in four studies—valuing conditions that they themselves were inflicted with—and from medical experts in six studies (Table 1). However, since the methods to recruit study participants included advertising the survey on a web site [14–26], a newspaper advertisement [27], recruitment in shopping centres [28] and train stations [29], e-mails to university students [16,18–22,30], and/or employees [31–33] a critical bias may have been introduced in the samples. Moreover, there was limited geographical diffusion of the studies, as thirty-one of the studies were conducted in North America and one in the Middle East.

Hence, there is a critical risk of bias in the selection of participant subdomains [11].

Regarding data completeness, a few of the studies had a high rate of non-responders (up to 43%), incompletely answered surveys, or exclusion of surveys due to illogical answers (Supplement 2). There is probably a serious risk of bias in the completeness of the data domain [11].

With regard to measurement instruments, SG, TTO, VAS, and WTP are considered generic instruments with acceptable validity and reliability, but there can still be a substantial risk for bias depending on how they are used [11]. All of the instruments used were constructed in-house, and only one study field tested and pre-assessed their survey for language, duration, and comprehension ($n = 30$) [34]. None of the studies provided detailed information on the reliability and validity of the instruments used. All but three studies [27–29] used computer-based administration of the instrument. Few studies have included visual prompts used in their publications. Hence, there are some unclaritys regarding the validity and reliability of the instruments [35]. Moreover, there are unclaritys regarding the representation of the studied health states [11]. For example, most of the studies included only one scenario, with one photo of the health state that was presented

Table 1. Preferences for health states treated with plastic surgery.

| First author, year, country | Whose references (perspective) | Health state ^a pre-treatment | Preferences before surgery mean (SD) | Health state post treatment | Preferences after surgery mean (SD) | Risk of bias |
|-------------------------------|---|---|--|---|---|--------------|
| Yeung, 2011, USA [56] | Med stud (n = 171) | General principles and procedures Flap surgery/ microsurgery | | Scar donor site from free flap harvest -anterolateral thigh -proximal lateral calf -lateral arm | VAS: 0.82 (0.19) VAS: 0.50 (0.20) VAS: 0.22 (0.16) | Serious |
| Chuback, 2015, Canada [29] | | Burns Severe face and neck burn and composite tissue allotransplantation after facial burn | | Composite tissue allotransplantation after facial burn | | Serious |
| | Gen pop (n = 25) | | TTO: 0.68 SG: 0.54 | | TTO: 0.82 SG: 0.78 | |
| | Med exp (n = 25) | | TTO: 0.42 SG: 0.40 | | TTO: 0.63 SG: 0.63 | |
| | Pat (n = 25) | | TTO: 0.57 SG: 0.54 | | TTO: 0.67 SG: 0.64 | |
| Hultman, 2015, USA [46] | Pat (n = 20) | Burn scars, non- specified location | | Burn scar dyschromia treated with intense pulsed light | WTP-C: 1200 (2664) | Serious |
| | | Head and neck Congenital Cleft lip and palate | | | | |
| Sinno, 2012, Canada, USA [16] | Gen pop (n = 104) Med stud (n = 6) | | VAS: 0.69 (0.18) TTO: 0.85 (0.16) SG: 0.84 (0.18) | | | Serious |
| Byun, 2016, Canada [17] | Gen pop (n = 104) | Microtia without deafness | VAS: 0.90 (0.85–0.95)* TTO: 0.91 (0.87–0.95)* SG: 0.95 (0.91–0.97)* | | | Serious |
| | | Microtia with deafness | VAS: 0.90 (0.72–0.85)* TTO: 0.88 (0.77–0.91)* SG 0.91 (0.84–0.97)* | | | |
| Kuta, 2017, Canada [30] | Gen pop (n = 75) Med stud (n = 43) | Scaphocephaly due to sagittal craniosynostosis | VAS: 0.85 (0.76–0.95)* TTO: 0.91 (0.84–0.95)* SG: 0.92 (0.85–0.98)* | | | Serious |
| | | Traumatic and neoplastic Severe face and neck burn and composite tissue allotransplantation after facial burn | | Composite tissue allotransplantation after facial burn | | Serious |
| Sinno, 2010, Canada, USA [18] | Gen pop (n = 142) Med stud (n = 114) | Facial disfigurement requiring facial transplantation | VAS: 0.46 (0.02) TTO: 0.68 (0.03) SG: 0.66 (0.03) | | | Serious |
| Dey, 2016, USA [15] | Gen pop (n = 200) | Facial deformity, without surgery <i>Small peripheral</i> | VAS: 0.76 (0.25) SG: 0.94 (0.14) WTP-C: \$1170 (4082) WTP-I: \$1213 (1580) | Facial deformity, after surgery <i>Small peripheral</i> | VAS: 0.95 (0.15) SG: 1.0 (0.01) WTP-C: \$50 (288) WTP-I: \$7 (54) | Moderate |
| | | <i>Small central</i> | VAS: 0.56 (0.26) SG: 0.83 (0.22) WTP-C: \$3240 (4653) WTP-I: \$4089 (3247) | <i>Small central</i> | VAS: 0.94 (0.16) SG: 0.99 (0.06) WTP-C: \$88 (350) WTP-I: \$35 (120) | |
| | | <i>Large peripheral</i> | VAS: 0.52 (0.27) SG: 0.81 (0.24) WTP-C: \$5125 (10 088) WTP-I: \$4572 (3595) | <i>Large peripheral</i> | VAS: 0.96 (0.12) SG: 1.0 (0.01) WTP-C: \$116 (1152) | |
| | | <i>Large central</i> | VAS: 0.38 (0.26) SG: 0.72 (0.26) WTP-C: \$7875 (15 892) | <i>Large central</i> | WTP-I: \$11 (69) VAS: 0.81 (0.26) SG: 0.94 (0.15) WTP-C: \$1785 (5265) | |

(continued)

Table 1. Continued.

| First author, year, country | Whose references (perspective) | Health state ^a pre-treatment | Preferences before surgery mean (SD) | Health state post treatment | Preferences after surgery mean (SD) | Risk of bias |
|-------------------------------|--|---|---|--|---|--------------|
| | | | WTP-I: \$9446 (6338) | | WTP-I: \$1330 (2745) | |
| Sinno, 2012, Canada, USA [40] | Gen pop (n = 96) Med stud (n = 5) | Facial palsy Unilateral facial paralysis | VAS: 0.56 (0.18) TTO: 0.78 (0.21) SG: 0.79 (0.21) | | | Serious |
| Su, 2017, USA [26] | Gen pop (n = 384) | Facial paralysis, grade <i>Normal</i> | VAS: 0.99 (0.18) TTO: 1.00 (0.03) SG: 1.00 (0.19) WTP-C***: \$1980 (1233–3014) | | | Moderate |
| | | <i>Low</i> | VAS: 0.92 (0.18) TTO: 0.98 (0.11) SG: 0.98 (0.09) WTP-C***: \$3487 (2363–4961) | | | |
| | | <i>Medium</i> | VAS: 0.76 (0.21) TTO: 0.92 (0.21) SG: 0.93 (0.15) WTP-C***: \$8571 (6401–11 234) | | | |
| | | <i>High</i> | VAS: 0.43 (0.21) TTO: 0.74 (0.30) SG: 0.77 (0.25) WTP-C***: \$20 431 (16 273–25 317) | | | |
| Farris, 2018, USA [34] | Gen pop (n = 298) | Facial paralysis without surgery <i>Flaccid unilateral</i> | VAS: 0.60 (0.21) TTO: 0.72 (0.25) SG: 0.71 (0.24) | Facial paralysis post-facial reanimation <i>All types</i> | VAS: 0.74 (0.19) TTO: 0.84 (0.19) SG: 0.83 (0.21) | Serious |
| | | <i>Unilateral moderate to severe post-paralytic facial nerve syndrome</i> | VAS: 0.63 (0.21) TTO: 0.74 (0.25) SG: 0.75 (0.24) | | | |
| Almadani, 2020, Canada [23] | Gen pop (n = 463) | Mandibular hypoplasia | VAS: 0.48 (0.24) | Mandibular hypoplasia post distraction osteogenesis | VAS: 0.63 (0.20) TTO: 0.83 (0.13) | Serious |
| | Gen pop (n = 463) | Maxillary hypoplasia | VAS: 0.56 (0.23) | Maxillary hypoplasia after Le Fort I distraction | VAS: 0.73 (0.22) TTO: 0.83 (0.14) | Serious |
| | | Chest and breast | | | | |
| | | Neoplastic | | | | |
| Sinno, 2012, Canada, USA [37] | Gen pop (n = 116) Med stud (n = 4) | Bilateral mastectomy defect | VAS: 0.70 (0.18) TTO: 0.85 (0.16) SG: 0.86 (0.17) | | | Serious |
| Sinno, 2014, Canada, USA [39] | Gen pop (n = 149) Med stud (n = 5) ^b | Unilateral mastectomy defect | VAS: 0.75 (0.17) TTO: 0.87 (0.14) SG: 0.86 (0.18) | | | Serious |
| Ibrahim, 2015, Canada [21] | Gen pop (n = 97) Med stud (n = 6) | Nipple-areolar-complex deformity | VAS: 0.84 (0.18) TTO: 0.92 (0.11) SG: 0.92 (0.11) | | | Serious |
| Thoma, 2003, Canada [45] | Med experts (n = 33) | | | Breast reconstruction with free TRAM flap- no complications | VAS: 0.87 | Serious |
| Thoma, 2004, Canada [43] | Med exp (n = 32) | | | Breast reconstruction with free TRAM/DIEP-no complications | VAS: 0.87 | Serious |
| | | Other | | | | |
| Kerrigan, 2000, USA [27] | Female gen pop (n = 34) | Breast hypertrophy | SG: 0.94 VAS: 0.85 SG: 0.96 VAS: 0.75 | | | Serious |
| Chang, 2001, USA [14] | Med stud (n = 13) Gen pop (n = 355) | Breast hypertrophy | | | | Serious |
| | | <i>Mild</i> | VAS: 0.93 (0.081) TTO: 1.0 (0.089) SG: 0.98 (0.11) | | | |
| | | <i>Severe</i> | VAS: 0.70 (0.17) TTO: 0.85 (0.24) | | | |

(continued)

Table 1. Continued.

| First author, year, country | Whose references (perspective) | Health state ^a pre-treatment | Preferences before surgery mean (SD) | Health state post treatment | Preferences after surgery mean (SD) | Risk of bias |
|---|--|---|---|--|---|---------------------|
| Ibrahim, 2015, Canada [20] | Gen pop (n = 97) Med stud (n = 10) | Breast ptosis | SG: 0.88 (0.19) VAS: 0.80 (0.14) TTO: 0.87 (0.18) SG: 0.90 (0.14) | | | Serious |
| Mercili, 2019, USA [57] | Med exp (n = 20) | Trunk and abdomen | | Reconstruction of oncologic spinal wound, no complications | VAS: 0.92 | Serious |
| Sinno, 2014, Canada, USA [36] | Gen pop (n = 111) Med stud (n = 3) | Lower extremity Lower extremity lymphoedema | VAS: 0.50 (0.18) TTO: 0.76 (0.22) SG: 0.76 (0.21) | | | Serious |
| Chung, 2011, USA [58] | Med exp (n = 65) | | | Lower limb salvage and reconstruction after open tibia fracture, no complications | SG: 0.98 | |
| Heiser, 2020, USA [31] | Gen pop (n = 262) | Skin and adnexa Facial hypertrophic capillary malformation | VAS: 0.71 (0.24) TTO: 0.83 (0.21) SG: 0.83 (0.23) | Laser treated facial hypertrophy capillary malformation | VAS: 0.87 (0.16) TTO: 0.92 (0.16) SG: 0.91 (0.18) | Serious |
| Sinno, 2012, Canada, USA [38] | Gen pop (n = 8) Med stud (n = 120) | Aesthetic surgery Head and neck Nasal deformity after primary rhinoplasty | VAS: 0.80 (0.14) TTO: 0.91 (0.12) SG: 0.91 (0.13) | | | Serious |
| Aldihan, 2019, Saudi Arabia [28] | Gen pop (n = 161) Med stud (n = 129) Pat (n = 117) | Unattractive nose with functional problem | VAS: 0.77 (0.13) TTO: 0.87 (0.16) SG: 0.91 (0.17) VAS: 0.81 (0.13) TTO: 0.89 (0.15) SG: 0.96 (37) VAS: 0.71 (0.12) TTO: 0.83 (0.08) SG: 0.86 (0.16) | | | Serious |
| Kumar, 2020, USA [24] | Gen pop (n = 161) | Unattractive nose | VAS: 0.68 (0.21) | Cosmetic rhinoplasty | VAS: 0.36 (0.31) ^{****} SG: 0.90 (0.21) WTP ^{***} : \$15 681 (13426.09) | Moderate |
| Chen, 2019, USA [25] Faris, 2019, USA [32] | Gen pop (n = 228) Gen pop (n = 273) | Unattractive nose Post-rhinectomy nasal defect, untreated | VAS: 0.59 (0.24) TTO: 0.74 (0.24) SG: 0.74 (0.24) | Cosmetic rhinoplasty Post-nasal surgical reconstruction Post-nasal prosthetic rehabilitation | WTP: 2254-6949 VAS: 0.88 (0.16) TTO: 0.89 (0.13) SG: 0.9 (0.18) VAS: 0.67 (0.22) TTO: 0.82 (0.2) SG: 0.80 (0.2) | Moderate Serious |
| Abt, 2018, USA [33] | Harvard stud (n = 308) | Hair loss without surgery <i>Female</i> <i>Male</i> | VAS: 0.83 (0.19) TTO: 0.91 (0.18) SG: 0.92 (0.17) VAS: 0.85 (0.18) TTO: 0.93 (0.17) SG: 0.93 (0.17) | Hair loss after transplantation <i>Both sexes</i> | VAS: 0.93 (0.11) TTO: 0.95 (0.16) SG: 0.95 (0.15) | Serious |
| Sinno, 2011, Canada, USA [41] | Gen pop (n = 42) Med stud (n = 40) ^c | Postbariatric surgery Massive weight loss requiring body contouring | VAS: 0.79 (0.13) TTO: 0.89 (0.12) SG: 0.89 (0.15) | | | Serious |
| Sinno, 2012, Canada, USA [42] | Gen pop (n = 88) Med stud (n = 14) | Aging neck following massive weight loss | VAS: 0.89 (0.07) TTO: 0.94 (0.08) SG: 0.95 (0.10) | | | Serious |
| Izadpanah, 2013, Canada, USA [22] | Gen pop (n = 103) Med stud (n = 9) | Thigh laxity after massive weight loss | VAS: 0.77 (0.15) TTO: 0.90 (0.11) SG: 0.89 (0.14) VAS: 0.80 (0.14) | | | Serious |

(continued)

Table 1. Continued.

| First author, year, country | Whose references (perspective) | Health state ^a pre-treatment | Preferences before surgery mean (SD) | Health state post treatment | Preferences after surgery mean (SD) | Risk of bias |
|---------------------------------|---|---|--------------------------------------|-----------------------------|-------------------------------------|--------------|
| Ibrahim, 2014, Canada, USA [19] | Gen pop (<i>n</i> = 97) Med stud (<i>n</i> = 10) | Arm laxity after massive weight loss | TTO: 0.91 (0.12) SG: 0.94 (0.10) | | | |

*Median (interquartile range).

**Mean (95%, CI).

***Willingness to pay for the reconstruction depicted in the study.

****Change after surgery.

^aDivided into domains according to the European Training Requirements for the Specialty of Plastic, Reconstructive, and Aesthetic Surgery [9].

^bdemography is given for 154 participants, even though only 140 unique answers were analysed.

^ceducation is only given for 82 participants.

DIEP: Deep inferior epigastric perforator flap; Gen pop: General public; Med exp: Medical expert; Med stud: Medical student; Pat: Patients with the condition in question; SG: Standard gamble; TRAM: Transverse rectus abdominis myocutaneous flap; TTO: Time trade-off; VAS: Visual analogue scale; WTP: Willingness to pay; WTP-C: Willingness to pay determined by the standard contingent valuation method; WTP-I: Willingness to pay derived from the iterative bidding method.

to the participants, which means that it cannot be representative of the health state of interest. Moreover, only a minority of the studies gave a brief statement on how they described the clinical scenarios, and not all of them showed photos of the cases that had been used. Therefore, it is unclear what decision the study population has been faced with. Most of the studies verified whether the participants understood the instrument (Supplement 2). The frequency of surveys excluded from analysis because subjects gave binocular blindness a higher score than monocular blindness, in a control scenario, varied from 3.3% to 26%. Hence, there is a critical risk for bias in the measurement instrument subdomain.

With regard to data analysis, none of the studies performed statistical analyses that were adjusted for potential confounders [11], although some of the studies compared differences between subgroups using Chi-square and Student's *t*-tests. As a result, there could be a serious risk of bias in the data analysis subdomain.

With regard to publication bias, all studies were conducted by plastic surgeons, which could be a sign of publication bias as surgeons are probably more inclined to report preferences that imply that surgery is necessary. All articles were published in surgical journals. Nonetheless, as this was an issue across all studies, the risk of bias arising from expertise was considered low. On the other hand, 38% of the studies (13/34) were conducted by the same research group (Sinno et al.) [16,18–22,36–42], four studies were conducted by Heiser et al. [31–34], and two by Dey et al. [15,26], which could have introduced some bias. None of the studies were industry-sponsored. In summary, the risk of publication bias was low.

When the risk of bias was weighed together across the different domains, it was considered moderate for four studies [15,24–26] and serious for the rest (Table 1). At the body of evidence level, the risk of bias was serious.

Indirectness

Most of the studies had PICO-related sources of indirectness (also called representativeness, generalisability, external validity, and applicability) [11], as the majority of the general population consisted of students, which are not necessarily representative of the general population and therefore not optimal. However, as the total evidence has been downgraded due to issues of bias in the samples, it was not further downgraded due to indirectness [11].

Inconsistency

There are very few plastic surgical health states for which more than a single study evidence exists (Table 1). Two studies assessed breast hypertrophy, three evaluated facial palsy, and five studied different forms of 'unattractive noses' (Table 1). However, these studies included very different descriptions of health states [12]. Assessment of consistency was difficult because none of the studies provided confidence intervals [12]. The results were too different to be pooled [12]. As a result, there is probably a serious risk of inconsistency across studies.

Imprecision

Most studies had a sample size of 100 respondents or more but studies conducted on patients and medical experts had fewer respondents (*n* = 20–117) (Table 1). According to the GRADE Working Group, a sample size of 380 people is needed to elicit preferences for choices that entail a choice with closely balanced benefits and harms (more than 55% of patients would make the same choice with an error margin of 55%) and 139–249 for a choice with potentially large benefits (80–90% of people would make the same choice with an error margin of 5%) [12]. None of the studies calculated confidence intervals for preferences (Table 1). Based on this, it is unclear whether the sample sizes were large enough to reduce the risk of chance. As the total evidence has been downgraded due to issues of bias in the samples, it was not further downgraded due to imprecision [12].

Results of individual studies and overall evidence

The lowest preference scores were generally elicited for facial defects/anomalies and the highest for excess skin after massive weight loss. In the head and neck studies, scores tended to be similar or higher for congenital malformations, such as cleft lip and palate (TTO 0.85), microtia (TTO 0.88–0.91), craniosynostosis (TTO 0.91), than for acquired defects, such as facial paralysis (TTO 0.72–0.98), and rhinectomy (TTO 0.74). In the breast studies, scores were similar for breast hypertrophy (TTO 0.85), breast ptosis (TTO 0.87), and mastectomy defects (TTO 0.85–0.87) (Table 1, Figure 2). Only one study has investigated the health states that arose after unsuccessful cosmetic surgery [38]. Scores for an unsuccessful rhinoplasty (TTO 0.91) were higher than scores for a congenital 'unattractive nose' (TTO 0.83) and the results of nasal reconstruction after rhinectomy due to cancer (TTO 0.89) (Table 1). Fourteen studies examined the health status after successful reconstructive surgery (Table 1). Studies comparing pre- and

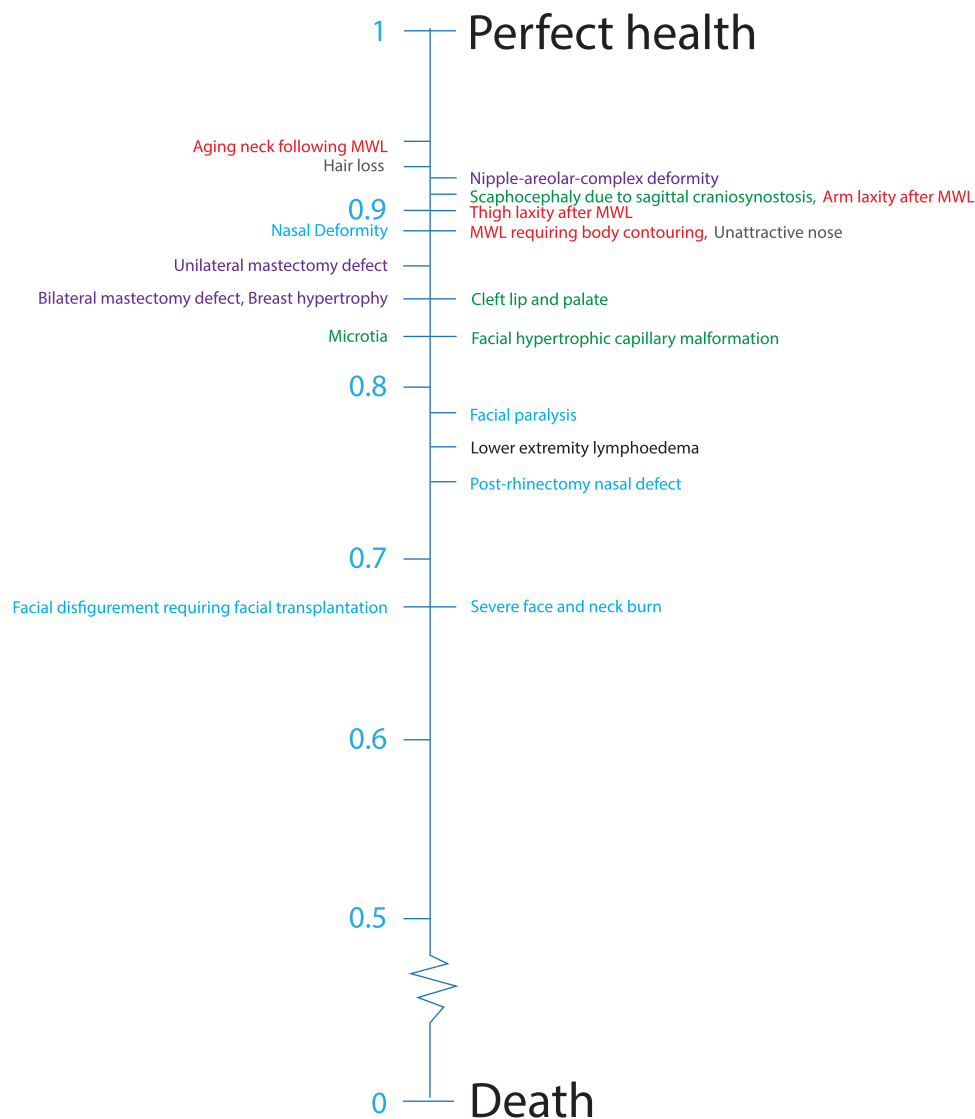


Figure 2. Preference measurements according to TTO as stated by the general public. Red: Excess skin after MWL (massive weight loss), Blue: Acquired facial deformities, Green: Congenital facial malformations, Purple: Breast deformities, Black: Lower extremity, Grey: Aesthetic defects. In cases where several scores exist for the same condition, the lowest score was used.

post-reconstructive results generally showed that the scores were higher after the operation. Several studies performed on breast reconstruction [43–45] revealed that scores are similar between autologous breast reconstruction without complications (TTO 0.85) as for mastectomy defects (TTO 0.85–0.87). Five studies reported WTP estimates for health states covering facial paralysis [26], facial deformity [15], unattractive nose [24,25], and burn scars [46], with estimates between \$1170 for treating a small peripheral facial deformity [15] to \$20,431 for a high grade of facial paralysis [26] (Table 1). Studies with a moderate risk of bias revealed that more pronounced facial deformity [15] and facial paralysis [26] generate lower scores than those with less severe deformities and that surgical correction elevates the scores considerably [15].

The overall body of evidence for preferences for health states treated with plastic surgery was low (GRADE $\Theta\Theta\Theta$). The evidence was rated due to a serious risk of bias and inconsistencies.

Discussion

In this systematic review, several studies measuring preferences for health states treated with plastic surgery using direct methods

were identified. However, the risk of bias was considered serious for most of them, and the body of evidence was generally low.

Methodological limitations of the included studies

Several methodological aspects must be considered when evaluating the studies. There was a risk of bias regarding the selection of participants in the studies, data completeness, measurement instruments, data analysis, and a risk of publication bias [11,12].

Valuation studies are affected by sampling and recruitment, that is which respondents provided preference scores [47], which introduces a considerable risk of bias in this review. There are different views on whether preferences should be elicited from patients, general public, or medical experts. These perspectives might constitute different indirectness problems depending on what the preferences are used for; for example, if they are used to make decisions in health care, to decide on health care policies, or to develop clinical guidelines [11]. It has been argued that preferences on plastic surgical health states, which generally affect appearance, should be elicited from the general public as patients tend to become accustomed to a certain defect and

therefore undervalue the effect of the health state [4,48]. It has also been pointed out that family members and caregivers of children with malformations might have the best understanding of what a health state entails but should not be used to elicit preferences as parents are generally unwilling to take any risk or gamble with their children which introduces a bias [4]. In the present review, most of the included studies extracted preferences from a sample of the general population (Table 1). Notwithstanding, the samples are relatively small (Table 1: median 104 participants (range 8–384)), introducing an imprecision, with a narrow demographic variation with regard to ethnicity, culture, age, and education level, and the majority of the studies were conducted in North America (Table 1). Hence, it is unclear whether the samples can be considered as general populations. In addition, many of the studies (13/34) were conducted by the same research group (Sinno et al.) [16,18–22,36–42], two studies by Ishii, LE et al. [15,26], and two by Heiser et al. [33,34] (Table 1), and there is no information in the articles regarding the different samples related to each other; that is, if a separate questionnaire was sent for each health state or several health states were tested in the same sample. In summary, it is uncertain whether the preference scores can be generalised.

The results could be affected by how the health states were described and depicted [49], which constitutes a major indirectness problem and a critical risk of bias in this review. Most plastic surgical health states comprise a wide spectrum of clinical conditions. For example, cleft lip and palate (CLP) covers everything from an isolated alveolar cleft, without functional consequences, bilateral total CLP with hearing impairment and speech difficulties, and syndromic clefts with considerable impact on different cognitive and physical functions [50]. In the preference study on CLP included in this study [16], only one scenario was presented to the participants. The same is valid for most of the health states examined in the included articles. Therefore, the elicited preference scores are probably not representative of the health states in large, but merely for the specific example presented. Another problem is that the results can be affected by what the health states are compared to and how the anchors—‘perfect health’ and ‘worst health’—are described [50]. Most of the included studies only investigated one health state and no healthy controls; therefore, possible inconsistencies in scores could not be examined. In most studies, monocular and binocular blindness was included as a control, and participants who scored binocular blindness higher, which is less of a handicap than monocular blindness, were excluded since the scoring was considered as an indicator of a poor understanding of the methodology. However, this method does not exclude other forms of inconsistencies in scoring that can occur when different health states are not compared to each other or to healthy controls. The valuation may also be affected by unrelated factors of the patient, such as age, race, and gender [50]. Such factors could have a more pronounced effect when health states involve an appearance-related component as most of the health states are treated with plastic surgery. Similarly, the labelling of health states may have an impact on valuation [50]. Most of the scenarios used in the included articles were labelled with diagnosis and how the condition arose, for example, ‘aging neck following massive weight loss’ [42] or ‘mastectomy defect’ [37,39]. Hence, it cannot be ruled out that factors such as prejudice and assumptions about certain conditions or treatments and what they entail, could have affected the scores. In short, the description and illustration of the scenarios used are of particular relevance when health states

treated by plastic surgery are investigated as they routinely involve appearance-related defects.

The validity and reliability of different techniques for measuring preferences vary, and different methods often lead to different preference values [51]. In general, the VAS method generates the lowest scores, while the SG method generates the greatest [51]. This can be clearly seen in the studies included in the present review, as the TTO and SG scores were fairly similar, whereas the VAS scores were considerably lower (Table 1). The preferences elicited by WTP consistently showed logical results, in that higher values were produced for postoperative than preoperative conditions. However, differences in methodology used and health states analysed in the different studies make it impossible to draw any conclusions.

Can preferences be used to prioritise in plastic surgery?

As we indicated in the introduction, the extent to which plastic surgery should be prioritised for public funding is a matter for debate, and a clear and consistent rationale remains elusive. In many healthcare jurisdictions, plastic surgery with no obvious relationship to biomedical functionality is a ‘borderline’ case when it comes to public funding—where some interventions are funded and some are not. In most cases, aspects such as the degree of need and cost-effectiveness play an important role in delimitating publicly funded plastic surgery [1]. Scientific knowledge about preferences and the resulting health gains (e.g. QALYs), in relation to both conditions and their respective interventions, is an essential input both to assess the degree of need and the extent to which an intervention will meet that need and be cost-effective. However, the outcome of preference studies might also play an essential role in deciding which plastic surgery procedures should be considered for public funding or rationed within the system, implying out-of-pocket spending. For example, if conditions or complications from interventions result in preferences in line with preferences for statistically normal conditions not receiving treatment, it might not warrant public funding [52].

A basic requisite for such applications is that the preference scores have been elicited in a scientifically sound way and that they are relevant to the health care system in which they are going to be used. In the review, we found that autologous breast reconstruction without complications [44] received the same score as that for a mastectomy defect [39]. Should this be interpreted as an indication that we might abstain from breast reconstruction and use these resources elsewhere in the healthcare system, or is it rather a measuring error? Even if there are studies indicating a similar and different quality of life over time between patients undergoing reconstruction [53,54], we might not be willing to take these results at face value as grounds for priority setting.

Moreover, this raises the need for an in-depth discussion on whose preferences should be elicited, the general public (i.e. the taxpayers) or affected patients. When it comes to delimit what should be publicly funded, there is a case of taking the general public’s perspective. WTP estimates are often difficult to use in a priority setting in a healthcare system striving for equity because of its strong correlation with income and ability to pay [55]. Indirect methods of eliciting preferences, often with the use of the EQ-5D instrument, are often used instead of direct methods (but not included in our systematic review). However, if the direct methods measuring preferences for health states treated with plastic surgery are poor, one may question the value of such indirect methods for the same health states.

Conclusions

The overall certainty of evidence for directly measured preference values and WTP for health states treated by plastic surgery is low (GRADE $\Theta\Theta\Theta\Theta$). However, scientific knowledge about preferences and the resulting health gains (e.g. QALYs) might play an essential role in deciding what type of plastic surgery should be considered for either public funding or rationed within the system. Better quality studies are required to allow for such applications.

Note

1. Minority populations <10 individuals have not been accounted for here. All details are given in Table 1.

Acknowledgements

We thank librarian Emelie Mellbrand, The Medical Library, Sahlgrenska University Hospital, Gothenburg, for their excellent help with information sources and search strategies, and medical photographers Niclas Löfgren and Johanna Hedborg, Department of Plastic and Reconstructive Surgery, Sahlgrenska University Hospital, Gothenburg for their skilful help with figures.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Emma Hansson  <http://orcid.org/0000-0002-3218-0881>
Lars Sandman  <http://orcid.org/0000-0003-0987-7653>
Thomas Davidson  <http://orcid.org/0000-0002-9393-2573>

References

- [1] Henderson J. The plastic surgery postcode lottery in England. *Int J Surg*. 2009;7(6):550–558.
- [2] Russell J, Swinglehurst D, Greenhalgh T. ‘Cosmetic boob jobs’ or evidence-based breast surgery: an interpretive policy analysis of the rationing of ‘low value’ treatments in the English National Health Service. *BMC Health Serv Res*. 2014;14:413
- [3] Cook SA, Rosser R, Meah S, et al. Clinical decision guidelines for NHS cosmetic surgery: analysis of current limitations and recommendations for future development. *Br J Plast Surg*. 2003;56(5):429–436.
- [4] Sinno H, Dionisopoulos T, Slavin SA, et al. The utility of outcome studies in plastic surgery. *Plast Reconstr Surg Glob Open*. 2014;2(7):e189
- [5] Thoma A, Haines T, Veltri K, et al. A methodological guide to performing a cost-utility study comparing surgical techniques. *Can J Plast Surg*. 2004;12(4):179–187.
- [6] Thoma A, McKnight LL. Quality-adjusted life-year as a surgical outcome measure: a primer for plastic surgeons. *Plast Reconstr Surg*. 2010;125(4):1279–1287.
- [7] Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. *Br Med Bull*. 2010;96:5–21.
- [8] Gustavsson E. Characterising needs in health care priority setting [PhD thesis]. Linköping: Linköping University; 2017.
- [9] Cooke A, Smith D, Booth A. Beyond PICO: the SPIDER tool for qualitative evidence synthesis. *Qual Health Res*. 2012;22(10):1435–1443.
- [10] UEMS. European Training Requirements for the Specialty of Plastic, Reconstructive and Aesthetic Surgery, European Standards of Postgraduate Medical Specialist Training. Brussels: European Union of Medical Specialists/Union Européenne des médecins spécialistes (UEMS); 2014.
- [11] Zhang Y, Alonso-Coello P, Guyatt GH, et al. GRADE guidelines: 19. Assessing the certainty of evidence in the importance of outcomes or values and preferences-Risk of bias and indirectness. *J Clin Epidemiol*. 2019;111:94–104.
- [12] Zhang Y, Coello PA, Guyatt GH, et al. GRADE guidelines: 20. Assessing the certainty of evidence in the importance of outcomes or values and preferences-inconsistency, imprecision, and other domains. *J Clin Epidemiol*. 2019;111:83–93.
- [13] Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924–926.
- [14] Chang WT, Collins ED, Kerrigan CL. An internet-based utility assessment of breast hypertrophy. *Plast Reconstr Surg*. 2001;108(2):370–377.
- [15] Dey JK, Ishii LE, Joseph AW, et al. The cost of facial deformity: a health utility and valuation study. *JAMA Facial Plast Surg*. 2016;18(4):241–249.
- [16] Sinno H, Tahiri Y, Thibaudeau S, et al. Cleft lip and palate: an objective measure outcome study. *Plast Reconstr Surg*. 2012;130(2):408–414.
- [17] Byun S, Hong P, Bezuhly M. Public perception of the burden of microtia. *J Craniofac Surg*. 2016;27(7):1665–1669.
- [18] Sinno HH, Thibaudeau S, Duggal A, et al. Utility scores for facial disfigurement requiring facial transplantation [outcomes article]. *Plast Reconstr Surg*. 2010;126(2):443–449.
- [19] Ibrahim AM, Sinno HH, Izadpanah A, et al. Population preferences of undergoing brachioplasty for arm laxity. *Ann Plast Surg*. 2014;73(Supplement 2):S149–S52.
- [20] Ibrahim AM, Sinno HH, Izadpanah A, et al. Mastopexy for breast ptosis: utility outcomes of population preferences. *Plast Surg (Oakv)*. 2015;23(2):103–107.
- [21] Ibrahim AM, Sinno HH, Izadpanah A, et al. Nipple-areolar complex reconstruction following postmastectomy breast reconstruction: a comparative utility assessment study. *Plast Reconstr Surg Glob Open*. 2015;3(4):e380.
- [22] Izadpanah A, Sinno H, Vorstenbosch J, et al. Thigh laxity after massive weight loss: a utilities outcomes assessment. *Ann Plast Surg*. 2013;71(3):304–307.
- [23] Almadani YH, Gilardino MS. Cost-effectiveness analysis, psychosocial, and utility outcomes of early mandibular distraction in craniofacial microsomia. *J Craniofac Surg*. 2020;31(7):1888–1894.
- [24] Kumar AR, Ishii M, Papel I, et al. The health utility and valuation of cosmetic rhinoplasty. *Facial Plast Surg Aesthet Med*. 2020;22(4):268–273.
- [25] Chen D, Ishii M, Nellis J, et al. Assessment of casual observers’ willingness to pay for increased attractiveness through rhinoplasty. *JAMA Facial Plast Surg*. 2019;21(1):27–31.
- [26] Su P, Ishii LE, Joseph A, et al. Societal value of surgery for facial reanimation. *JAMA Facial Plast Surg*. 2017;19(2):139–146.
- [27] Kerrigan CL, Collins ED, Kneeland TS, et al. Measuring health state preferences in women with breast hypertrophy. *Plast Reconstr Surg*. 2000;106(2):280–288.
- [28] Aldihan K, Alnasyan A, Albassam A, et al. Comparing the health burden of living with nasal deformity in actual patients and healthy individuals: a utility outcomes score assessment. *Ann Plast Surg*. 2019;83(4):381–383.

- [29] Chuback J, Yarascavitch B, Yarascavitch A, et al. Measuring utilities of severe facial disfigurement and composite tissue allotransplantation of the face in patients with severe face and neck burns from the perspectives of the general public, medical experts and patients. *Burns*. 2015;41(7):1524–1531.
- [30] Kuta V, McNeely PD, Walling S, et al. Sagittal craniosynostosis: a utility outcomes study. *J Neurosurg Pediatr*. 2017;20(2):113–118.
- [31] Heiser A, Jowett N, Occhiogrosso J, et al. Societal-perceived health utility of hypertrophic facial port-wine stain and laser treatment. *Facial Plast Surg Aesthet Med*. 2020;22(5):327–335.
- [32] Faris C, Heiser A, Quatela O, et al. Health utility of rhinectomy, surgical nasal reconstruction, and prosthetic rehabilitation. *Laryngoscope*. 2020;130(7):1674–1679.
- [33] Abt NB, Quatela O, Heiser A, et al. Association of hair loss with health utility measurements before and after hair transplant surgery in men and women. *JAMA Facial Plast Surg*. 2018;20(6):495–500.
- [34] Faris C, Tessler O, Heiser A, et al. Evaluation of societal health utility of facial palsy and facial reanimation. *JAMA Facial Plast Surg*. 2018;20(6):480–487.
- [35] Brazier J, Deverill M. A checklist for judging preference-based measures of health related quality of life: learning from psychometrics. *Health Econ*. 1999;8(1):41–51.
- [36] Sinno H, Izadpanah A, Tahiri Y, et al. The impact of living with severe lower extremity lymphedema: a utility outcomes score assessment. *Ann Plast Surg*. 2014;73(2):210–214.
- [37] Sinno H, Izadpanah A, Thibaudeau S, et al. An objective assessment of the perceived quality of life of living with bilateral mastectomy defect. *Breast*. 2013;22(2):168–172.
- [38] Sinno H, Izadpanah A, Thibaudeau S, et al. The impact of living with a functional and aesthetic nasal deformity after primary rhinoplasty: a utility outcomes score assessment. *Ann Plast Surg*. 2012;69(4):431–434.
- [39] Sinno H, Izadpanah A, Vorstenbosch J, et al. Living with a unilateral mastectomy defect: a utility assessment and outcomes study. *J Reconstr Microsurg*. 2014;30(5):313–318.
- [40] Sinno H, Thibaudeau S, Izadpanah A, et al. Utility outcome scores for unilateral facial paralysis. *Ann Plast Surg*. 2012;69(4):435–438.
- [41] Sinno H, Thibaudeau S, Tahiri Y, et al. Utility assessment of body contouring after massive weight loss. *Aesth Plast Surg*. 2011;35(5):724–730.
- [42] Sinno HH, Ibrahim AM, Izadpanah A, et al. Utility outcome assessment of the aging neck following massive weight loss. *Otolaryngol Head Neck Surg*. 2012;147(1):26–32.
- [43] Thoma A, Veltri K, Khuthaila D, et al. Comparison of the deep inferior epigastric perforator flap and free transverse rectus abdominis myocutaneous flap in postmastectomy reconstruction: a cost-effectiveness analysis. *Plast Reconstr Surg*. 2004;113(6):1650–1661.
- [44] Offodile AC, 2nd, Chatterjee A, Vallejo S, et al. A cost-utility analysis of the use of preoperative computed tomographic angiography in abdomen-based perforator flap breast reconstruction. *Plast Reconstr Surg*. 2015;135(4):662e–669e.
- [45] Thoma A, Khuthaila D, Rockwell G, et al. Cost-utility analysis comparing free and pedicled TRAM flap for breast reconstruction. *Microsurgery*. 2003;23(4):287–295.
- [46] Hultman CS, Friedstat JS, Edkins RE. Efficacy of intense pulsed light for the treatment of burn scar dyschromias: a pilot study to assess patient satisfaction, safety, and willingness to pay. *Ann Plast Surg*. 2015;74(Suppl 4):S204–S208.
- [47] Boyd NF, Sutherland HJ, Heasman KZ, et al. Whose utilities for decision analysis? *Med Decis Making*. 1990;10(1):58–67.
- [48] Barker JH, Furr A, Cunningham M, et al. Investigation of risk acceptance in facial transplantation. *Plast Reconstr Surg*. 2006;118(3):663–670.
- [49] Stalmeier PF, Goldstein MK, Holmes AM, et al. What should be reported in a methods section on utility assessment? *Med Decis Making*. 2001;21(3):200–207.
- [50] Crockett DJ, Goudy SL. Cleft lip and palate. *Facial Plast Surg Clin North Am*. 2014;22(4):573–586.
- [51] Doctor JN, Bleichrodt H, Lin HJ. Health utility bias: a systematic review and meta-analytic evaluation. *Med Decis Making*. 2010;30(1):58–67.
- [52] Sandman L, Hansson E. An ethics analysis of the rationale for publicly funded plastic surgery. *BMC Med Ethics*. 2020 2;21(1):94.
- [53] Parker PA, Youssef A, Walker S, et al. Short-term and long-term psychosocial adjustment and quality of life in women undergoing different surgical procedures for breast cancer. *Ann Surg Oncol*. 2007;14(11):3078–3089.
- [54] Harcourt DM, Rumsey NJ, Ambler NR, et al. The psychological effect of mastectomy with or without breast reconstruction: a prospective, multicenter study. *Plast Reconstr Surg*. 2003;111(3):1060–1068.
- [55] Widstrom E, Seppala T. Willingness and ability to pay for unexpected dental expenses by Finnish adults. *BMC Oral Health*. 2012;12:35.
- [56] Yeung LC, Ellstrom CL, Martin MC. A donor-site preference utility study for three flaps used in lower extremity microvascular reconstruction. *Ann Plast Surg*. 2011;66(1):59–61.
- [57] Mericli AF, Rhines L, Bird J, et al. Immediate reconstruction of oncologic spinal wounds is cost-effective compared with conventional primary wound closure. *Plast Reconstr Surg*. 2019;144(5):1182–1195.
- [58] Chung KC, Shauver MJ, Saddawi-Konefka D, et al. A decision analysis of amputation versus reconstruction for severe open tibial fracture from the physician and patient perspectives. *Ann Plast Surg*. 2011;66(2):185–191.