Eliciting preferences for clinical follow-up in head and neck cancer patients using best-worst scaling

Michela Meregaglia1,2, MSc, John Cairns1,3, PhD, Salvatore Alfieri4, MD, Federica Favales4, MSc, Daniela Mazzitelli5, MSc, Ester Orlandi6, MD, Lisa Licitra4, MD, Paolo Bossi4, MD

1 Department of Health Services Research and Policy; Faculty of Public Health and Policy; London School of Hygiene and Tropical Medicine, London (UK)

2 CeRGAS (Research Centre on Health and Social Care Management), Bocconi University, Milan (Italy)

3 CCBIO, University of Bergen, Bergen (Norway)

4 Head and Neck Medical Oncology Department, IRCCS Foundation National Cancer Institute, Milan (Italy)

5 School of Short Integrated Psychotherapy, ISeRDiP, Milan (Italy)

6 Radiation Therapy Department, IRCCS Foundation National Cancer Institute, Milan (Italy)

Corresponding author:

Michela Meregaglia, Via Roentgen 1, 20136 Milan (Italy)

Tel. +39-02-5836-5935. Email: michela.meregaglia@unibocconi.it.

Abstract

Objectives: There are no commonly accepted standards for monitoring patients treated for head and neck cancer. The aim of this study was to assess patients' preferences for different aspects of follow-up.

Methods: A best-worst survey was conducted in a sample of head and neck cancer patients in clinical follow-up at the National Cancer Institute (Milan, Italy). Conditional logit regression with choice as the dependent variable was run to analyse the data. A covariate-adjusted analysis was performed in order to

identify socio-demographic and clinical factors related to the selection of best-worst items. The participants were asked to report any difficulties encountered during the survey.

Results: A total of 143 patients, predominantly male (74%) and with a mean age of 58 years were enrolled in the survey. The strongest positive preference was expressed for a hospital-based program of physical examinations with frequency decreasing over time. Conversely, the lowest valued item was not performing any positron emission tomography (PET) scan during follow-up. Patients with high educational levels were more likely to value attending a primary care-based program and undergoing intensive radiological investigations. Other patient-specific variables significantly associated with the choice of items were employment and living status, time already spent in follow-up and number of treatments received.

Conclusions: Overall, patients were more likely to choose an intensive follow-up scheme broadly consistent with the program currently administered by the hospital. There is little evidence of preference heterogeneity that might justify customized programs based on demographics. The best-worst scaling task appeared feasible for most participants.

Introduction

Head and neck cancer (HNC) is the sixth most common cancer in the world with nearly 700,000 new diagnoses and 370,000 deaths reported each year [1,2]; incidence in Italy has been assessed at 7.7 cases per 100,000 [3]. HNC consists of a heterogeneous group of malignancies affecting several anatomical sites and with different prognoses [4]. The main risk factors are alcohol and tobacco abuse and, in recent years, the infection with human papilloma virus (HPV) for cancers located in the oropharynx [1]. The risk of relapse is greater in the first two years after primary treatment, when an estimated 50-60% of patients develop loco-regional recurrences or metastases [1]; lifetime risk of second primary cancers is around 10-20% (i.e. 2%-5% per year) [5]. Therefore, a follow-up program is essential shortly after the completion of treatment in order to identify potentially curable relapses. However, the optimal timing of visits and radiological assessments following treatment is debated by oncologists. Published recommendations are mostly informed by retrospective studies, expert opinions and clinical practice rather than trial-based evidence [2,5]. Until now, no consensus has been reached on the optimal follow-up modalities and timing in HNC patients.

In addition to this clinical uncertainty, the patient's perspective has traditionally been neglected in designing cancer programs and elaborating clinical guidelines, whilst considering individual preferences might improve the feasibility, acceptability and effectiveness of healthcare interventions [6]. This is part of a larger study evaluating the cost-effectiveness of different surveillance schemes in HNC [HETeCo, clinicaltrials.gov identifier NCT02262221]. The current objective is to quantify preferences for post-treatment surveillance in a large sample of patients treated for primary HNC.

Methods

This study used best-worst scaling (BWS) to elicit patient's preferences for different aspects of follow-up after primary treatment for HNC.

Experimental design

The BWS choice experiment is a variant of the widely adopted binary choice experiment approach. As in traditional discrete choice experiments (DCEs), this method requires the identification of key characteristics (i.e. the attributes) each of which is split into two or more levels, in order to create a series of scenarios described by different attribute-level combinations. However, instead of selecting

one scenario in a choice set of two or more, participants are asked to indicate which attribute-level they consider to be the 'best' and which to be the 'worst' (i.e. the BW pair) within each scenario. In other words, they choose 'the pair that exhibits the largest perceptual difference on an underlying continuum of interest' [7,8]. There exist three types of BWS studies in the literature: the object case (case 1), the profile or attribute case (case 2), and the multi-profile case (case 3) [9]. The present study used the profile case in which participants are presented with a series of different scenarios to be evaluated one at a time.

Our analysis was limited to the process-related aspects of the follow-up [10], as clinical outcomes of post-treatment surveillance in HNC are still under debate in the scientific community. Relevant attributes and levels were established from literature review and expert opinion. We searched common databases (PubMed and EMBASE) using key terms such as "cancer" AND "follow up" AND "discrete choice experiment" (OR "best worst") in title/abstract in order to identify studies that assessed patient's preferences around post-treatment programs in oncology using stated preference methods. Interviews with six patients during routine hospital visits were used to refine terminology and evaluate the comprehension and the acceptability of the BWS instrument.

After this preliminary work, we eventually identified four attributes: frequency and setting (hospital or mixed with primary care) of physical investigations; frequency of radiological assessments (magnetic resonance imaging - MRI or computed tomography - CT); frequency (and eligibility) of positron emission tomography (PET) scans; telephone calls to monitor the occurrence of new symptoms. Levels were presented in order of increasing intensity of care and resources consumption for each attribute (Table 1). A balanced study design was adopted in which each study attribute (K=4) had the same number of levels (LK = 3). If an alternative contains K attributes, there are K (K-1) = 4 (4-1) = 12 possible BW pairs the participant can choose within each scenario. As a full factorial design generating all possible attribute-level combinations (34 = 81 scenarios) was not feasible, a subset of 9 orthogonal scenarios (fractional factorial, main-effects design) was derived using the Hann and Shapiro catalogue, Master Plan 3 [11]. The total number of BW pairs in the orthogonal design was 108 (12·9). This sub-group of selected scenarios preserved the properties of orthogonality (i.e. each attribute-level appears an equal number of time in combination with all other attribute-levels) and balance (i.e. each level within an attribute appears an equal number of times) [12,13].

Recruitment and setting

Patients aged 18 years and over, with a diagnosis of HNC in any anatomical site (except for the skin) in the last 5 years, who had completed any curative treatment at the National Cancer Institute (NCI) in Milan were eligible to participate. Patients were excluded if they were unable to comply with the study in the opinion of the clinical investigators, or they could not provide their informed consent. Moreover, we excluded patients who underwent minor surgery for early stage cancer and subsequently did not attend a regular follow-up program in a multidisciplinary setting, i.e. with the contemporary presence of the head and neck surgeon, the radiation oncologist, and the medical oncologist. At NCI, the routine follow-up program consists of outpatient visits every 2-3 months for the first 2 years after the end of treatment, then every 5-6 months for 3 more years. Radiological evaluations with MRI/CT scan are performed once 3 months after the end of treatment, then annually. PET is requested only in the case of doubtful imaging; no scheduled inter-visit contact is planned during the follow-up period.

The study was described to a consecutive sample of eligible patients during a routine follow-up appointment. Patients were reassured that responses to the questionnaire would not affect the care they were receiving at the hospital [14]. Those who agreed to participate were asked to sign a consent form and received the survey. Socio-demographic and clinical information were collected for each study

participant. The questionnaire included a short rationale for the study and an explanation of the task required. The study was approved by the NCI Ethical Committee in March 2015; the survey was subsequently administered between May and October of the same year.

Statistical analysis

Data on patients' characteristics were summarized through descriptive statistics; categorical variables were presented as percentages while continuous variables were presented as means and standard deviations. In regression analyses, missing demographic data were imputed using logical rules and information from related variables or, whenever this approach was not feasible, the most common value (i.e. the mode) [15]. Missing BW responses were imputed with the items most frequently selected as best and worst respectively within each scenario. The number of times each item was chosen as 'best' or 'worst' by the study participants was calculated. A best-minus-worst score was calculated by subtracting the number of times a feature was chosen as worst from the number of times it was chosen as best [6,14].

Regression analysis was performed using a conditional logit model (clogit command in Stata) with cluster-adjusted (robust) standard errors [16]. BW pairs were treated as single variables and plotted as one data point at the individual level [17]. For each possible pair, the attribute-level was coded as 1 for the best and -1 for the worst; all remaining attribute-levels were coded as 0. The dependent variable took the value of 1 for the BW pair selected and 0 otherwise. In order to avoid a saturated model, the item that showed the lowest utility was used as reference level; the omitted item took the value of zero on the utility scale and all estimates of the model were interpreted in relation to that. Therefore, each attribute-level can be positioned on an underlying preference scale $(0; +\infty)$ starting with the reference item [9,18]. Statistically significant coefficients indicated the importance of the attribute-level in determining overall utility [19].

A covariate-adjusted analysis was also performed to investigate sub-group preferences according to socio-demographic and clinical characteristics. A conditional logit model was run as previously described; however, interaction factors between selected covariates and choice outcomes (i.e. attribute-levels) were also added as independent variables [16]. In this model, interaction coefficients represent the additional utility of each attribute-level for the covariate [6]. A preliminary univariate regression analysis was performed to identify the demographic variables to be included in the final covariate-adjusted model as those displaying significant interaction terms (p<0.05). Variables with three (or more) categories in the questionnaire were dichotomized to increase the sample size within each group. Two age classes were generated around its median value (59 years) [6]. With regards to the clinical variables, the number of treatments received (i.e. one vs. more than one) was chosen as a 'proxy' of disease severity that, according to the clinicians involved in the study, might influence patient's preferences in follow-up; the time from the end of treatments was equally considered clinically relevant. Any other clinical information was disregarded in this analysis.

All data were analyzed using Stata version 14 (Stata Corp, 2015).

Results

Sample characteristics

A total of 162 consecutive patients who met the inclusion criteria were approached to participate in the survey; however, sixteen declined resulting in a response rate of 90%. Three questionnaires were excluded from data analysis, as they were not completed correctly or in full. Therefore, the final sample comprised 143 patients, of whom 74% were male. Socio-demographic and clinical features of the

participants are presented in table 2. The mean age of participants was 57.6 (±12.1) years and more than one third of patients were retired (34.2%). The great majority of patients (85.3%) lived with family and 64.3% less than 100 km from the hospital.

A variety of primary tumor diagnoses were observed in the sample, with the most common being oropharyngeal (38.4%), nasopharyngeal (28.0%) and laryngeal cancer (11.2%), mostly in a locally advanced stage (III and IV; 93.7%). Most patients (38.5%) received a combination of chemotherapy and radiation as primary treatment for HNC, or chemotherapy followed by the combined therapy (30.0%). Participants were equally distributed according to time since the end of treatments as follows: ≤2 years, 51.0%; >2 years, 49.0%.

BWS frequency counts

Frequency counts provide summary estimates of best and worst choices made by participants (Table 3). Of a total of 2,574 expected BW responses, only 12 (0.5%) were missing and imputed as previously explained. The highest ranked attribute-level was "physical investigations performed every 2-3 months for 2 years, then every 5-6 months for 3 more years". The lowest rated feature is less clearly identifiable. According to the best-minus-worst score, the lowest valued attribute-level was "inter-visit calls by the nurse" to monitor patient's health status. "No PET scan during follow-up" was the item least frequently chosen as "best", while "primary care-based follow-up during the last 2 years" was that most often indicated as "worst".

Conditional logistic regression analysis

The logistic regression results are presented in table 4. The attribute-level with the lowest utility coefficient was "not performing any PET scan during follow-up" and was assumed as the reference level. The regression coefficients of BW pairs show the additional utility of each attribute-level over the reference case. As already observed in frequency counts, the feature showing the highest utility was "physical investigations performed every 2-3 months for 2 years and every 5-6 months for 3 more years". A more intensive frequency of visits ("every 2-3 months for 5 years") ranked second, and "MRI/CT scan performed once or twice a year" ranked third. In contrast, the attribute-levels with the lowest utility were "follow-up based at primary care during the last 2 years", "inter-visit calls by the nurse" and "no inter-visit calls from the hospital" in that order; however, none of them was statistically significant compared to the reference level.

For each individual attribute, the distance between the most and the least preferred levels is an indication of the relative importance of that attribute to respondents [9,18]. In this survey, the "frequency and setting of physical and larynx/pharynx endoscopic investigations" is the item with the largest difference between level coefficients (2.482 i.e. 2.523 minus 0.041) and, thus, the greatest impact on patients' utilities. Summing the level coefficients taken one at a time within each attribute, it is possible to calculate an overall utility for each hypothetical follow-up scheme deriving from the experiment. The most preferred scenario (overall utility: 6.120) across the sample would be a hospital-based follow-up with frequency of visits decreasing over time (i.e. every 2-3 months for 2 years, every 5-6 months for the next 3 years), radiological assessments (i.e. MRI/CT) performed once or twice a year, yearly PET scan for all patients (irrespective of individual risk of recurrences) and inter-visit calls by the oncologist to monitor the occurrence of new symptoms. On the contrary, the least desirable option is a mixed hospital-/primary care-based surveillance with MRI/CT scan performed only at the occurrence of new symptoms, no PET scan scheduled during follow-up period and inter-visit calls by the nurse to check the patient's health (scenario utility: 0.654).

Covariate-adjusted regression analysis

Table 5 provides results from the conditional logistic regression analysis after adjusting for selected clinical and demographic data. Educational level (more educated i.e. university, post-university =1; less educated i.e. primary school, secondary school =0), employment status (employed i.e. full-time employed, part-time employed, self-employed =1; not employed i.e. retired, unemployed, other =0), living status (with family=1; alone =0), time in follow-up (>2 years =1; ≤2 years =0) and number of treatments (≥1 i.e. 2, 3 or 4 =1; one only =0) which displayed significant interactions in univariate regression analysis (Tables S1-S8) were included in the final model. Conversely, no significant interaction coefficients were found with respect to age (age≥59 =1; age<59 =0), gender (female =1; male =0) and distance from home (≥100 km =1; <100 km =0).

The interpretation of regression results is facilitated through the example of education. There were statistically significant differences between education groups with respect to three of the four attributes in the experiment. The total utility of "MRI/CT scan performed once or twice a year" for highly educated patients is the sum of the attribute-level coefficient (1.590) and its interaction term coefficient with educational level (0.756), which gives 2.346. The corresponding utility for less educated people is the coefficient without interaction (1.590). Thus, we can infer that all patients like a more intensive radiological investigations program; however, this preference is stronger for those with more education. Furthermore, highly educated patients are more likely to prefer a primary care based follow-up than those with a lower education level. The total utility of this item, indeed, is equal to 1.878 for the former, while not significantly different from the reference value for the latter (0.178). The last significant interaction is with "no PET scan during follow-up"; the overall utility for more educated patients is 0.201 whilst not significantly different from zero for the less educated ones.

In a similar way, it is possible to calculate separate utilities for different groups of patients according to the remaining four covariates selected within the univariate analysis.

Patients' evaluation of the experiment

Table 6 presents data on patients' self-reported difficulties in understanding and completing the questionnaire. The average compilation time was 9.2 (± 3.1) minutes. Nearly half of the participants rated the BWS task very easy to perform (i.e. level 1; 45.4%) and did not need any support from healthcare professionals or family members (44.7%). More than one-third (37.1%) reported no difficulties during completion among the options available; "understanding the task" was the most common difficulty (21.7%) followed by "length of the questionnaire" (6.3%) and "technical/scientific language" (5.6%). A further 14% indicated other difficulties mainly related to indecision in selecting the BW pair and the feeling that scenarios were too repetitive.

Discussion

Few studies have explored patient's preferences for delivery of post-treatment cancer programs and even fewer have attempted to derive utility estimates from them. A non-systematic literature review identified two binary DCE studies [20,21] exploring women's preferences for breast cancer follow-up services in the Netherlands and Australia, respectively. Face-to-face contacts were strongly preferred to telephone ones and a more intensive program of visits (every 3-6 months) was preferred over the less intensive options. Moreover, women liked to be followed-up by a medical specialist and at specialized breast cancer clinics. A BWS study of post-treatment surveillance for soft tissue sarcoma in the UK concluded that patients typically preferred visits routinely consisting of a clinical examination and a chest X-ray, and secondary care- rather than general practice-based programs [17]. A further study [6] adopting the BWS methodology was identified even if not strictly related to follow-up but addressing a symptom supporting care intervention in lung cancer patients after completion of first line therapies.

With respect to HNC, a non-DCE survey only on patients' view of their follow-up regimen was conducted in UK. The study revealed that most patients felt their follow-up visits too frequent and were in favor of a less-intensive, symptom-driven follow-up [22].

The BWS method is argued to have several advantages over traditional DCEs [6,18]. First, respondents are provided with profiles one by one rather than two (or more) at a time; thus, BWS is considered less cognitively demanding for participants [8,16,19,23]. These expectations were confirmed in our study by the self-reported judgement on the choice task, which was graded as simple and quick by the majority of respondents. Moreover, BWS may elicit more information than traditional DCEs, as respondents make choices within profiles rather than between profiles; in particular, in BWS a single attribute-level combination acts as benchmark, instead of a whole scenario. In this way, it is possible to calculate utility coefficients for each item in the experiment, which may be useful in evaluating different elements of a healthcare service [8,16,19,23]. Profile-based (case 2) BWS was selected in preference to the traditional pair-wise DCE because it was anticipated that patients would always select the option which they thought would maximize survival and consequently less information would be generated by responses to the pair-wise choice task.

This study is the first stated preference survey of HNC follow-up and, in Italy, of any cancer surveillance. The survey aimed at providing insights into patients' views on post-treatment monitoring in this cancer population using BWS methodology. Moreover, a covariate-adjusted analysis was performed to investigate socio-demographic or clinical characteristics related to the choice of attribute-levels. It was not surprising to find that patients' preferences for HNC follow-up were generally aligned with the scheme currently adopted by NCI where the study was conducted. This tendency has been described as the 'lure of the familiar' [17], meaning that individuals are likely to stick with they have already experienced, even if potentially unsatisfactory. Participants in this study revealed clear preferences for follow-up to remain in secondary care, even during the last phases of the program. Intensive radiological examinations (once or twice a year) were strongly preferred. Inter-visit telephone calls were generally disliked, especially when performed by healthcare professionals other than medical doctors. These results are in contrast with those found by a previous study on patients' preferences in HNC follow-up [22]; however, that survey was conducted in a different geographical setting (highly deprived areas of London) and without relying on stated preference methods. Differences in preferences according to individual characteristics were also found. Overall, highly educated patients were more likely to prefer primary-care based follow-up and intensive MRI/CT radiological investigations but avoiding PET scan. Patients with a job tended to prefer more frequent visits to the hospital but no inter-visit calls, while those living with family revealed a stronger preference for not performing any PET scan during follow-up. Patients two years (or more) following treatment expressed a lower utility for more intensive MRI/CT investigations and were keener to avoid inter-visit telephone calls with their clinicians. Patients who had received more than one treatment option (e.g. surgery followed by radiotherapy) were less keen to accept a symptom-driven radiological surveillance and to travel frequently to the hospital for physical investigations. Conversely to a previous study [6], we found no differences in age or gender with regards to preferences for delivering a post-treatment intervention in cancer care; however, the program under evaluation was considerably different.

This study has a number of limitations. First, the data collection was restricted to only one center that, due to some distinctive features (i.e. high specialization, commitment to research, cutting-edge technologies), may not be representative of a typical cancer clinic in Italy. Moreover, patients attending NCI, especially those coming from afar, are likely to be more educated, wealthy or health conscious than the general HNC population. However, the referral to a single center reduced bias related to

different ways of administering the survey and providing support during the completion of the questionnaire. Second, the cognitive ability of each participant with respect to completing the task was not evaluated and in some occasions the patient was supported by the accompanying person. Nevertheless, given the very low number (n=3) of questionnaires excluded from the analysis and the limited self-reported difficulties, we are confident that the task was feasible for most participants. The final limitations concern the restricted range of hypothetical follow-up programs that can be valued owing to the small number of attributes-levels and the assumption of no interaction between BWS items. However, the limited number of items included in the experiment, as well as the use of a small factorial main-effects design (i.e. 9 out of 81 scenarios), was justified by feasibility considerations.

In recent years, there has been growing interest in using preference elicitation methods to inform health policy and medical decision-making. Incorporating patient's preferences into the treatment and follow-up strategies may help in tailoring healthcare to the patient and increase adherence to treatment [24]. In HNC follow-up, patients seem to be reassured by a regular follow-up with scheduled imaging and expertise of specialists, as already reported in other experiences [25]. The present study highlighted patients' limited interest in alternative ways of delivering post-treatment services, such as symptom-driven surveillance, telephone monitoring or non-specialist follow-up. Healthcare professionals (e.g. general practitioners or nurses) other than specialist doctors were probably considered not skilled enough to conduct cancer follow-up. There might be a resistance to change from established to new types of service without adequate reassurance from the clinicians. In particular, patients with less education may benefit least from a patient-initiated follow-up owing to difficulties in understanding medical instructions. Overall, there is a need for improved communication for cancer patients to evaluate consciously the post-treatment phase and to promote self-managed symptoms monitoring [26]. Patients likely prefer intensive radiological assessment because of fear of disease recurrence; however, tests should be performed for clinical reasons and not (only) for patient's reassurance. The long-term effects associated with frequent and prolonged radiological scans should also be considered. In this regard, more efforts should be spent in order to identify the most cost-effective follow-up scheme in HNC, thus providing the scientific community and patients with evidence-based programs. A randomized trial comparing health and economic outcomes in this setting is ongoing [HETeCo, clinicaltrials.gov identifier NCT02262221]. Finally, differences in preferences were found according to the intensity of treatments received and the time already spent in follow-up; these results might justify a provision of different surveillance schemes based on these clinical variables, as already suggested by guidelines in the field [27]. Similarly, inter-visit calls appear to be more valuable in the initial phases of the follow-up than in the final ones, when patients may feel more confident of beating cancer.

Overall, this study provided useful insights into individual preferences for several aspects of post-treatment surveillance in HNC in Italy. Additional elements might be explored in the future, such as the level of scientific evidence, co-payment for extra-investigations and late side effects of intensive investigations. Currently, there is evidence of heterogeneity in preferences with respect to a limited number of patient's characteristics. More research also considering the costs of different follow-up regimens is required to justify the provision of customized follow-up programs in HNC patients.

References

[1] de Andrade DA, Machiels JP. Treatment options for patients with recurrent or metastatic squamous cell carcinoma of the head and neck, who progress after platinum-based chemotherapy. Curr Opin Oncol 2012; 24(3): 211-217.

[2] Kawecki A, Krajewski R. Follow-up in patients treated for head and neck cancer. Memo 2014; 7: 87-91.

[3] Wissinger E, Griebsch I, Lungershausen J et al. The economic burden of head and neck cancer: a systematic literature review. Pharmacoeconomics DOI 10.1007/s40273-014-0169-3.

[4] Gatta G, Botta L, Sanchez MJ, et al. Prognoses and improvement for head and neck cancers diagnosed in Europe in early 2000s: the EUROCARE-5 population-based study. European Journal of Cancer 2015; 51: 2130-2143.

[5] Manikantan K, Dwivedi RC, Sayed SI, et al. Current concepts of surveillance and its significance in head and neck cancer. Ann R Coll Surg Engl 2011; 93: 576-582.

[6] Molassiotis A, Emsley R, Ashcroft D, et al. Applying best-worst scaling methodology to establish delivery preferences of a symptom supportive care intervention in patients with lung cancer. Lung Cancer 2012; 77: 199-2014.

[7] Flynn A, Louviere JJ. Determining the appropriate response to evidence of public concern: the case of food safety. J Public Policy Mark 1992; 11:12-25.

[8] Flynn TN, Louviere JJ, Peters TJ, et al. Best-worst scaling: what it can do for healthcare research and how to do it. J Health Econ 2007; 26: 171-189.

[9] Marti J. A best-worst scaling survey of adolescents' level of concern for health and non-health consequences of smoking. Social Science and Medicine 2012; 75: 97-97.

[10] Benning TM, Kimman ML, Dirksen CD, et al. Combining individual-level discrete choice experiment estimates and costs to inform health care management decisions about customized care: the case of follow-up strategies after breast cancer treatment. Value Health 2012; 15: 680-689.

[11] Kocur G, Adler T, Hyman W, et al. Guide to forecasting travel demand with direct utility assessment. United States Department of Transportation, Urban Mass Transportation Administration, Report UMTA-NH-11-0001-82-1., Washington DC.

[12] Johnson FR, Lancsar E, Marshall D, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR conjoint analysis experimental design good research practices task force. Value Health 2013; 16: 3-13.

[13] Ungar WJ, Hadioonzadeh A, Najafzadeh M, et al. Quantifying preferences for asthma control in parents and adolescents using best-worst scaling. Respiratory Medicine 2014; 108: 842-851.

[14] Ejaz A, Spolverato G, Bridges JF, et al. Choosing a cancer surgeon: analysing factors in patient decision making using a best-worst scaling methodology. Ann Surg Oncol 2014; 21: 3732-3738.

[15] Duhamel A, Nuttens MC, Devos P, et al. A preprocessing method for improving data mining techniques. Application to a large medical diabetes database. Stud Health Technol Inform 2003; 95: 269-274.

[16] Flynn TN, Louviere JJ, Peters TJ et al. Estimating preferences for a dermatology consultation using best-worst scaling: comparison of various methods of analysis. BMC Medical Research Methodology 2008; 8:76.

[17] Damery S, Biswas M, Billingham L, et al. Patient preferences for clinical follow-up after primary treatment for soft tissue sarcoma: a cross-sectional survey and discrete choice experiment. Eur J Surg Oncol 2014; 1-7.

[18] Severin F, Schmidtke J, Muhlbacher A, Rogowski WH. Eliciting preferences for priority setting in genetic testing: a pilot study comparing best-worst scaling and discrete-choice experiments. European Journal of Human Genetics 2013; 21: 1202-1208.

[19] Franco MR, Howard KH, Sherrington C, et al. Eliciting older people's preferences for exercise programs: a best-worst scaling choice experiment. Journal of Physiotherapy 2015; 61: 34-41.

[20] Bessen T, Chen G, Street J, et al. What sort of follow-up services would Australian breast cancer survivors prefer if we could no longer offer long-term specialist-based care? A discrete choice experiment. BJC 2014; 1-9.

[21] Kimman ML, Dellaert BGC, Boersma LJ, et al. Follow-up after treatment for breast cancer: one strategy fits all? An investigation of patient preferences using a discrete choice experiment. Acta Oncologica 2010; 49: 328-337.

[22] Trinidade A, Kothari P, Andreou Z. Follow-up in head and neck cancer: patients' perspective. Int J Health Care Qual Assur 2012; 25(2); 145-9.

[23] Potoglou D, Burge P, Flynn T, et al. Best-worst scaling vs. discrete choice experiments: an empirical comparison using social care data. Social Science and Medicine 2011; 72: 1717-1727.

[24] de Souza JA, Seiwert TY. A value framework in head and neck cancer care. Am Soc Clin Oncol Educ Book 2014; e304-9.

[25] Lewis RA, Neal R, Hendry M, et al. Patients' and healthcare professionals' views of cancer follow-up: systematic review. Br J Gen Pract 2009; DOI: 10.3399/bjgp09X453576.

[26] Frew G, Smith A, Zutshi B, et al. Results of a quantitative survey to explore both perceptions of the purposes of follow-up and preferences for methods of follow-up delivery among service users, primary care practitioners and specialist clinicians after cancer treatment. Clinical Oncology 2010; 22: 874-884.

[27] De Felice F, Musio D, Tombolini V. Follow-up in head and neck cancer: a management dilemma. Advances in Otolaryngology 2015; http://dx.doi.org/10.1155/2015/703450.