Using stated preference methods to inform public health decision making: Evidence from Croatia

Dukić Samaržija, Nikolina

Source / Izvornik: **Management : Journal of Contemporary Management Issues, 2021, 26, 23 - 41**

Journal article, Published version
Rad u časopisu, Objavljena verzija rada (izdavačev PDF)

https://doi.org/10.30924/mjcmi.26.2.2

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:192:579500

USING STATED PREFERENCE METHODS TO INFORM PUBLIC HEALTH DECISION MAKING: EVIDENCE FROM CROATIA*

*Nikolina Dukić Samaržija***

Received: 4. 11. 2020
Accepted: 23. 11. 2021
DOI https://doi.org/10.30924/mjcmi.26.2.2

Preliminary communication
UDC 330.163:614

Abstract

Croatia has recognised the importance of prevention programmes in the field of public health, although their effectiveness is not satisfactory due to the low population response, which has a negative impact on the rationalisation of public spending. One of the possible solutions is to consider the stated preferences of the target population. Therefore, the main objective of this study is to evaluate the use of stated preference methods in improving public health prevention programmes. For the purpose of the study, a questionnaire was designed using three different methods - Discrete Choice Experiment (DCE), Contingent Rating method (CR) and Best-Worst Scaling method (BWS). This study shows that the attributes of the Croatian cervical cancer screening programme are significantly associated with the respondents' utility level, which in turn is related to women's response. Since BWS, DCE, and CR measure the same construct - utility - we can say that convergent validity partially confirms the external validity of the methods. The author concludes that it is necessary to implement market principles, i.e. the demand-side analysis using stated preference methods, in the planning, implementation and re-evaluation of public health programmes.

Keywords: *public health care, choice-based methods, prevention, decision making, Best-Worst Scaling*

1. INTRODUCTION

Chronic disease prevention (primary and secondary) has become one of the major public health issues in Europe (Espina et al., 2018). Effective public health prevention programmes reduce the need (American Cancer Society, 2018) and cost of medical treatments (Laviana et al., 2020; Wei-Hua et al., 2010; Sassi & Hurst, 2008), the latter being extremely important in times of crisis and reduced budgetary spending. The target population of public health prevention programmes are healthy individuals who have not developed symptoms of a disease. Therefore, their perception of prevention benefits, which is reflected in the possibility of reducing the risk of disease in the future, is reduced. When making decisions about public health programmes, decision-makers should consider the utility derived from process attributes, i.e. it is necessary to consider non-medical

* This work has been fully supported by the University of Rijeka under the project number [uniri-društv-18-166].

** Nikolina Dukić Samaržija, Assistant professor, University of Rijeka, Faculty of Economics and Business, Ivana Filipovića 4, 51000 Rijeka, Croatia, Tel: +385 51 355 184, E-mail: nikolina.dukic.samarzija@efri.hr, ORCID: https://orcid.org/0000-0002-5158-5809

and procedural attributes as well as health outcomes (Trapero-Bertran et al., 2019).

It is believed that there is a strong link between the preferences and behaviour of economic agents, which is reflected in their choices within the available options (alternatives). Thus, by analysing the choices of individuals, one can identify their preferences and derive a utility function (Bockstael & McConnell, 2007). The author argues that simply by increasing the capacity on the supply side (gynaecologists, consulting rooms) and neglecting factors on the demand side (patients' preferences), it is impossible to achieve the desired effectiveness of public health prevention programmes (Dukić et al., 2015). Therefore, the main purpose of this research is to present the stated preference elicitation as a useful tool to improve public health prevention programmes as it can be used to improve the effectiveness of these programmes.

2. LITERATURE REVIEW

A growing body of research has used the discrete choice experiment (DCE) to elicit preferences from respondents (patients, payers, and commissioners) without asking them directly about their preferred options (Soekhai et al., 2019; Guttman et al., 2009). In DCE, options differ according to their properties (attributes) and individuals value options depending on the attribute values (levels). Louviere and Woodworth (1983) first applied the DCE method on subjects to whom they had presented different scenarios (profiles) composed of characteristics of hypothetical products that had been carefully designed following the rules of experimental design. To increase the amount of data generated by the DCE method, the best-worst scaling (BWS) method was developed (Finn & Louviere, 1992). In BWS, the respondents are asked to mark the most and least desirable attribute level proposed in the scenario. In this way, the respondents not only rate different combinations of attributes, but also provide information on their preferences regarding different attribute levels. Since the respondents now choose not only between the proposed scenarios but also between the proposed attribute levels, given a sufficient number of respondents, it is possible to estimate the average utility of a given attribute level (Marley et al., 2008). Since 2001, there has been a growing body of empirical research using discrete choice methods to elicit preference regarding health-related attributes. De Bekker-Grob et al. (2012) and Clark et al. (2014) published a systematic review on the use of choice-based methods in health care or health economics in general. Both papers reviewed studies on preferences of consumers, patients and health professionals for all types of health care resources.

Previous empirical studies on women's cervical cancer screening preferences were conducted with different objectives and using different research techniques and data analysis methods. Prior to our research, studies had been conducted to test women's knowledge, attitudes and preferences regarding prevention programmes (Wordsworth et al., 2006; Johnston et al., 1996; Nicoll et al., 1991); to establish the need for widespread dissemination of information to women (McCaffery & Irwig, 2005); to identify the reasons for the lack of response to prevention programmes (Cerigo et al., 2013; Kivistik et al., 2011; Fort et al., 2011; Holroyd et al., 2004); and to identify the determinants that influence the demand for prevention and prevention activities among women (Matejic et al., 2011; Holroyd et al., 2004). In contrast to the above studies, this study focuses on

the attributes of the prevention programme itself, namely the Croatian Cervical Cancer Screening Programme. Our study revealed there is a growing interest for DCE in analysing women preferences for cervical cancer screening attributes addressing different issues – from improving screening uptake (Li et al., 2019; Oberlin et al., 2019; Subramanian et al., 2018) to reducing screening inequalities (Raginel et al., 2020).

3. ANALYTICAL AND METHODOLOGICAL FRAMEWORK

The DCE method, the Contingent Rating method (CR), and the BWS method were used to design the questionnaire and analyse the responses. The actual process of empirical research was conducted in three main phases. In the first phase of the research, an analysis was conducted based on a systematic review of the current literature and consultation with public health experts, which led to the selection of programme attributes and their level. In the second phase of the study, the main survey was administered and data were collected. In the third phase, the collected data were analysed and the validity of the econometric models was assessed.

3.1. First phase: Selection of attributes and their levels

To better clarify the logic behind the first phase of the research, the process of creating hypothetical scenarios (hypothetical programmes) will be explained in more detail. First, the statistical properties have to be met (Louviere et al., 2000): avoidance of correlation between attributes, the balance of the experimental design, and minimal overlap. Apart from the statistical properties, cognitive complexity and market reality (Louviere et al., 2000), which are two essential non-statistical characteristics, have a significant impact on the design. Therefore, it is crucial to identify the attributes relevant for decision making and create attribute levels that are significant and realistic for the analysis. To this end, field research was conducted in the form of an interview with gynaecologists in primary health care, coordinators of the National Programme in the Institute of Public Health and women already invited for screening. This led to the final attributes and corresponding levels, the significance of which was confirmed by a pilot test on a sample of 64 women aged 25 to 64 years (Table 1).

Table 1. The attributes and corresponding levels

Characteristics (attributes)	Attribute levels	B[a]
Manner of scheduling examination – A1	by phone**	0.184
	in person	-1.300*
	online	1.119*
Waiting period for examination – A2	up to 2 weeks	0.609*
	from 2 weeks to 1 month	-0.36*
	from 1 to 3 months**	-0.248

Waiting time in the waiting room[b] – A3	up to 30 min	1.723*
	from 30 min to 1 hour	0.457*
	from 1 to 2 hours**	-2.180
Waiting period for findings – A4	up to 10 days	0.965*
	from 10 days to 1 month	0.179
	from 1 to 3 months**	-1.144
Manner of delivery of findings – A5	to the home address**	0.237
	phone call of the gynaecologist/nurse	0.592*
	personal arrival to the gynaecologist's office	-0.83*
Examination cost – A6	free**	2.198
	HRK 95 (cca 50% of the total cost[c])	-0.75*
	HRK 188 (100% of the total cost[c])	-1.45*

[a] Coefficients of the conditional logit model on the pilot study data.

[b] The pilot study was conducted for the following levels: the exact scheduled time, up to 30 minutes, from 30 minutes to 1 hour .

[c] Cost specification (Croatian Institute for Health Insurance): the Papanicolaou (Pap) test – HRK 100; gynaecological examination – HRK 45; cytological processing – HRK 43.

* $p < 0.05$

** Reference level attributes.

In order to obtain a balanced experimental design, i.e. the appearance of different attribute levels approximately the same number of times, through which the variance in the estimated parameters is minimised (Mangham et al., 2009), all the attributes have been described through three different levels.

3.2. Second phase: Questionnaire design and main survey implementation

In the second phase, the statistical experimental design of the questionnaire, hypothetical scenarios (different combinations of previously selected attribute levels) were created through the fractional factorial design of the main effects (Hall et al., 2001). In this research phase, two specific versions of the questionnaire (versions A and B) were created with nine different selection procedures (Louviere & Lancsar, 2009; Adamowitcz et al., 1998). For this study, a less cognitively demanding form of the BWS method (i.e. the profile case-BWS) was used (Flynn, 2010) because the study includes a population of women of different age groups, educational levels, geographic origins, and social and economic background. In addition to determining the preferences of the target population, this study aimed to verify the usefulness of the BWS method, which seems to be the best choice for this type of research. To demonstrate the advantages of the BWS method in relation to the contingent rating method (CR), the respondents were asked to rate the proposed hypothetical programmes (scenarios) using the Likert scale from 1 to 5 (1 if the respondent was very dissatisfied and 5 if the respondent was very satisfied with the hypothetical programme). In addition, with respect to the hypothetical programmes, the respondents were asked to indicate whether they would be willing to respond to screening, allowing the results of the BWS method to be compared with the results of

the binary logit model of the DCE method. Finally, by introducing different stated preference methods (BWS, DCE and CR) into the analysis, it was possible to assess the external validity of the BWS method in terms of the convergence of results. An example of a hypothetical programme from the questionnaire is shown below.

Table 2. An example of a hypothetical programme

Worst attribute *(only one)*	**Programme**	**Best attribute** *(only one)*
	You make an appointment ***by telephone***	
	You wait for an examination **from *1 to 3 months***	
	You wait for an examination in the waiting room ***from 1 to 2 hours***	
	You wait for findings ***from 10 days to 1 month***	
	You are informed of the findings **by telephone** by the gynaecologist/nurse	
	You pay for an examination ***HRK 95***	
On a scale (1-5), rate your satisfaction with the above-described programme.		
1 **2** **3** **4** **5**		
very dissatisfied		*very satisfied*
Are you willing to respond to the examination?		**YES** **NO**

Following the example of previous empirical studies, the survey also required the respondents to answer questions that provided information about their personal characteristics (age, education level, employment status, etc.).

3.3. Third phase: Statistical evaluation

The study used a simple binary discrete choice to compare the BWS method with the DCE method, while a binary logit model (de Bakker-Grob, 2009; McIntosh, 2003) was used for data analysis. The model was based on yes/no (0,1) responses of the respondents, which are also the simplest responses for using and interpreting the results obtained (Kjar, 2005). The respondents' answers (S = 9) were analysed by 2,313 observations (9×257). In addition, since these yes/no responses have a relatively low discriminatory power, a more precise format was introduced (the CV method). Ordinal logistic regression was used to specify the CV method.

In the case of BWS, paired data analysis at the respondent level was used as a part of the econometric evaluation of the model. Therefore, the reported choices of the best and worst attribute levels were the furthest on the latent utility scale. In other words, the respondents considered the best and worst attribute levels in the context of a single proposed hypothetical programme $K\times(K\text{-}1)$ possible pairs, which in this case would mean 30 (6×5) best-worst (BW) pairs for six attributes (K) according to the

presented scenario (programme). Since the questionnaire proposed nine scenarios (programmes), the respondents were confronted with 270 different BW pairs during the questionnaire, resulting in 270 observations per respondent. The amount of information obtained per questionnaire is one of the fundamental advantages of the BWS method compared to other preference elicitation methods.

The variables of the attribute impact (the manner of scheduling the examination, waiting period for the examination, waiting time in the waiting room, waiting period for the findings, manner of delivery of findings, and examination cost) are coded so that the attribute *K* influence ponder takes the value 1 for all BW pairs where attribute *K* was chosen best and the value -1 for all BW pairs where attribute *K* was chosen worst. The levels of attribute *K* (scale value) took the value 1 if they occurred as the best level in the BW pair, -1 otherwise, and 0 if the observed level was not proposed in the hypothetical scenario. Using this method of variable coding (effect coding), the values of the omitted attribute levels can be calculated by multiplying the sum of the remaining attribute levels by -1 (Flynn et al., 2008). Finally, the dependent variable (BW pairs) took the value of 1 for the selected BW pair from the proposed programme and 0 otherwise.

4. RESEARCH FINDINGS

4.1. Reference sample

Women in the 25 to 64 age group in the Primorje-Gorski Kotar County were included in the empirical study as only the women in this age group had been invited for screening organised by the National Health Institutes at the county level. A stratified statistical sample by the geographic origin was created from a population of 87,444 women aged 25 to 64 years (Health Statistical Yearbook of the Primorje-Gorski Kotar County, 2012). The sample was divided into four localities: Rijeka and its surroundings, the Liburnia area, Gorski Kotar, and the islands. The survey was conducted from December 2013 to April 2014. A total of 367 questionnaires was received (a response rate of 61%), out of which 67 questionnaires did not answer all the questions for the proposed hypothetical programmes, so they were excluded from the analysis. Finally, 300 (*N*) questionnaires were evaluated, meeting the criterion of the minimum required sample size of 300 respondents. Although the time period of the survey is no longer current, it was nonetheless conducted to elicit women's preferences for screening attributes in order to emphasise the patient-centred approach to public health interventions. The survey results are still relevant today, as the surveyed programme has been suspended and is still in the pilot phase.

In accordance with Orme's (2006) recommendations to use a minimum sample of 300 respondents for a robust data analysis, 600 questionnaires were given. Out of the 600 questionnaires, 375 (63%) were delivered to community health centres in Rijeka and the surrounding area, 98 (16%) to sites in Liburnia, 80 (13%) on the islands of Rab, Cres, Krk, and Lošinj, and 47 (8%) to a site in Gorski Kotar. The description of survey return according to geographical origin is presented in the following table.

Table 3. Survey return according to respondents' geographical origin

LOCATION	FREQUENCIES	PERCENTAGE
Rijeka and the surrounding area	146	48.67
The Liburnia area	74	24.67
Islands	57	19.00
Gorski Kotar	23	7.67
Total	300	100.00

It is evident that the distribution of the sample by region corresponds to the actual distribution of the population in the county.

4.2. The binary logistic regression results for the discrete choice experiment

The results of the binary logistic regression model are shown in Table 4.

Table 4. The results of the binary logistic regression

Attribute Levels	B	S.D.	Wald	df	Sig.
A1 - by phone*	-	-	-	-	-
A1 - in person	**-0.294**	0.204	2.077	1	0.150
A1 - online	**-0.057**	0.200	0.082	1	0.775
A2 - up to 2 weeks	**0.599**	0.201	8.853	1	0.003
A2 - from 2 weeks to 1 month	**0.551**	0.208	7.005	1	0.008
A2 - from 1 to 3 months*	-	-	-	-	-
A3 - up to 30 min	**0.398**	0.208	3.679	1	0.055
A3 - from 30 min to 1 hour	**0.273**	0.212	1.668	1	0.197
A3 - from 1 to 2 hours*	-	-	-	-	-
A4 - up to 10 days	**0.402**	0.225	3.191	1	0.074
A4 - from 10 days to 1 month	**0.251**	0.206	1.480	1	0.224
A4 - from 1 to 3 months*	-	-	-	-	-
A5 - to the home address*	-	-	-	-	-
A5 - phone call	**-0.301**	0.216	1.941	1	0.164
A5 - personal arrival	**-0.733**	0.199	13.509	1	0.000
A6 – free*	-	-	-	-	-
A6 - 95 HRK	**-1.126**	0.214	27.551	1	0.000
A6 - 188 HRK	**-2.350**	0.227	106.869	1	0.000

* Reference scenario

The model is based on the x^2 test ($H(^2)$ = 1498.798), significant at $p < 0.01$, while the Hosmer-Lemeshow test (p = 0.806) confirms that the model is well fitted to the data. The model parameters (B) are interpreted as influencing the likelihood of the programme being adopted by the respondents, depending on the marginal change in certain attribute levels. If the change in attribute levels is associated with an increased likelihood of programme adoption, an individual is assumed to have positive preferences with respect to the observed changes in attribute levels within the framework of random utility theory (Bridges, 2003).

Although all the estimated coefficients of the observed attribute levels did not achieve the significance level $p < 0.05$, it can be concluded that with respect to the proposed reference scenario, there was a significant correlation between the decision to participate in screening and the following attribute levels of the programme: waiting period for the examination (coefficient (utility) increases as waiting time decreases), manner of delivery of findings (appearing in person at the physician's office is negatively correlated with participation in screening), and examination cost (increase in the examination cost is negatively correlated with participation in screening). The explanation for attribute levels that were not significant can be found in the poorer discriminatory power of the selection method defined in this way (yes/no), which also indicates that a more precise format, such as the contingent rating method, should be used.

4.3. The ordinal logistic regression model results for the contingent rating

An ordinal logistic regression model was used for the economic specification and analysis of the data obtained by the contingent rating method, the results of which are presented in Table 5.

Table 5. A comparison of ordinal logistic regression results

ATTRIBUTE LEVELS	WITHOUT RATIONALITY TEST[a]					RATIONALITY TEST[b]				
	B	S.D.	Wald	df	Sig.	B	S.D.	Wald	df	Sig.
A1 - by phone*	-	-	-	-	-	-	-	-	-	-
A1 - in person	**-0.378**	0.105	12.957	1	0.000	**-0.758**	0.126	36.114	1	0.000
A1 - online	**-0.174**	0.105	2.759	1	0.097	**-0.324**	0.125	6.699	1	0.010
A2 - up to 2 weeks	**0.611**	0.106	33.537	1	0.000	**1.066**	0.128	69.525	1	0.000
A2 - from 2 weeks to 1 month	**0.200**	0.110	3.334	1	0.068	**0.430**	0.131	10.760	1	0.001
A2 - from 1 to 3 months*	-	-	-	-	-	-	-	-	-	-
A3 - up to 30 min	**0.664**	0.111	35.755	1	0.000	**1.208**	0.135	80.125	1	0.000
A3 - from 30 min to 1 hour	**0.377**	0.108	12.155	1	0,000	**0.665**	0.129	26.632	1	0.000
A3 - from 1 to 2 hours*	-	-	-	-	-	-	-	-	-	-
A4 - up to 10 days	**0.659**	0.109	36.575	1	0.000	**0.791**	0.130	36.760	1	0.000
A4 - from 10 days to 1 month	**0.303**	0.111	7.486	1	0.006	**0.565**	0.132	18.342	1	0.000
A4 - from 1 to 3 months*	-	-	-	-	-	-	-	-	-	-

A5 - to the home address*	-	-	-	-	-	-	-	-	-	-
A5 - phone call	**0.030**	0.108	0.076	1	0.783	**-0.064**	0.128	0.246	1	0.620
A5 - personal arrival	**-0.384**	0.105	13.436	1	0.000	**-0.626**	0.126	24.794	1	0.000
A6 – free*	-	-	-	-	-	-	-	-	-	-
A6 - 95 HRK	**-0.860**	0.106	65.226	1	0.000	**-1.131**	0.129	77.484	1	0.000
A6 - 188 HRK	**-1.333**	0.109	149.608	1	0.000	**-1.806**	0.133	183.923	1	0.000

[a] N = 300
[b] N = 244
*Reference scenario

The model has a statistically significant predictive power ($p < 0.001$), and there are no significant differences between the original and expected frequencies ($p > 0.05$). In addition, the percentage of outcome variance explained by the model is at a level that exceeds 47% (rx2 = 0.472). Furthermore, with respect to the reference scenario, which most closely corresponds to actual conditions, statistical significance was found at the significance level exceeding 95% for all attribute levels except online appointment, waiting period for examination from 2 weeks to 1 month, and telephone notification of findings.

The results of the ordinal logistic regression model indicate a correlation between programme attributes and the respondents' level of satisfaction with the hypothesised programmes. To empirically test this hypothesis, a nonparametric Kruskal-Wallis (K-W) test was conducted to test the correlation between response to the hypothesised surveys and satisfaction level as measured by a 1-5 Likert scale. The K-W test confirmed that the satisfaction level can explain the response to the hypothetical screening (H(2) = 480.266, $p = 0.00$). The mean ranks from 1 to 5 (424.84; 602.67; 1038.25; 1220.90; 1239.76) showed that the response increased with an increase in the respondents' satisfaction (utility) level.

Since the analysis of the questionnaires revealed a significant number of rated hypothetical scenarios that can be considered inconsistent with rational decision making, the rationality test followed (Louviere et al., 2000). The results of the rationality test have indicated some improvements in terms of the significance of the estimated coefficients (as shown in Table 5) and the percentage of outcome variance explained by the model, which is now more than 54% (rx2=0.543), while the model maintained its good predictive power ($p < 0.001$). Looking at the significant influences of the attribute levels on the expected log-transformed result of higher respondent satisfaction, it can be concluded that the following attribute levels are related to the decrease in the expected satisfaction level of the respondents (cf. Table 5): making an appointment for medical examination in person (-0.758), making an appointment for medical examination online (-0.324), obtaining information about the findings when arriving in person at the gynaecological clinic (-0.626), paying for the examination in the amount of HRK 95 (-1.131), paying for the examination in the amount of HRK 188 (-1.806). On the other hand, the attributes associated with the increase in the expected satisfaction level of the respondents are (cf. Table 5): waiting for examination up to 2 weeks (1.066), waiting for examination from 2

weeks to 1 month (0.430), waiting in the waiting room up to 30 minutes (1.208), waiting in the waiting room from 30 minutes to 1 hour (0.665), waiting for findings up to 10 days (0.791), waiting for findings from 10 days to 1 month (0.565).

4.4. The conditional logit model results for the best-worst scaling

The results obtained with the BWS method are presented in the following table, which shows the basic advantage of the BWS method over the traditional DCE method - the possibility of separating the attribute impact from the scale value.

Table 6. The results of the conditional logit model

	B	**SE**	**Wald**	**df**	**Sig.**
Attribute impact					
Manner of scheduling a medical examination – A1	**0.606**	0.049	155.855	1	0.000
Waiting period for examination – A2	**0.043**	0.048	0.816	1	0.366
Waiting time in the waiting room – A3	**0.399**	0.048	67.769	1	0.000
Period of waiting for the findings – A4	**0.225**	0.048	21.909	1	0.000
Manner of delivery of the findings – A5	**0.596**	0.049	150.623	1	0.000
Examination cost – A6	-	-	-	-	-
SCALE VALUE					
A1-1 by phone	**-0.058**	-	-	-	-
A1-2 in person	**-0.671**	0.068	97.735	1	0.000
A1-3 on-line	**0.729**	0.068	115.941	1	0.000
A2-from 1 to 2 weeks	**0.691**	0.071	93.380	1	0.000
A2-2 from 2 weeks to 1 month	**-0.301**	0.073	16.835	1	0.000
A2-3 from 1 to 3 months	**-0.390**	-	-	-	-
A3-1 up to 30 min	0.787	0.068	133.016	1	0.000
A3-2 from 30 min to 1 hour	**0.116**	0.072	2.590	1	0.108
A3-3 from 1 to 2 hours	**-0.904**	-	-	-	-
A4-1 up to 10 days	**0.993**	0.066	227.017	1	0.000
A4-2 from 10 days to 1 month	**-0.157**	0.071	4.872	1	0.027
A4-3 from 1 to 3 months	**-0.836**	-	-	-	-
A5-1 to the home address	**-0.407**	-	-	-	-
A5-2 by phone	**0.605**	0.070	75.360	1	0.000
A5-3 personal arrival	**-0.198**	0.070	7.880	1	0.005
A6-1 free	**1.887**	-	-	-	-
A6-2 95 HRK	-0.589	0.068	75.577	1	0.000
A6-3 188 HRK	**-1.299**	0.056	528.533	1	0.000

With respect to the reference attribute A6 (*examination cost*), the attributes A1 (*manner of scheduling the examination*) and A5 (*manner of delivery of findings*) had the greatest impact, while the parameter of the attribute A2 (*waiting for the examination*)

was not statistically significant, i.e. at $p < 0.05$, the attribute of waiting for the examination was not significantly different from the attribute of examination costs ($p = 0.36$). The model parameters (B) reflect the marginal utility associated with the attributes and attribute levels (Flynn et al., 2008). For an attribute with a negative parameter sign, the probability that the observed attribute (if *ceteris paribus*) is chosen best decreases as the parameter strength increases, while the probability that the same attribute is chosen worst increases (McFadden, 1973). To facilitate the visibility of attribute levels on the utility scale, Table 7 shows the order of attribute levels by their position on the utility scale.

Table 7. The positions of attribute levels on the utility scale

ATTRIBUTE LEVELS	SCALE VALUE	ATTRIBUTE LEVELS	SCALE VALUE
A6-1 free	1.887	A1-1 by phone	-0.058
A4-1 up to 10 days	0.993	A4-2 from 10 days to 1 month	-0.157
A3-1 up to 30 min	0.787	A5-3 personal arrival	-0.198
A1-3 on-line	0.729	A2-2 from 2 weeks to 1 month	-0.301
A2-from 1 to 2 weeks	0.691	A2-3 from1 to 3 weeks	-0.390
A5-2 by phone	0.605	A5-1 to the home address	-0.407
A3-2 from 30 min to do 1 hour	0.116	A6-2 95 HRK	-0.589
		A1-2 in person	-0.671
		A4-3 from 1 to 3 months	-0.836
		A3-3 from 1 to 2 hours	-0.904
		A6-3 188 HRK	-1.299

Source: Table 6

By using the BWS method, the impact of attributes can be separated from the scale value of its levels (Flynn et al., 2008) - although the "manner of delivery of the findings" is a relatively highly valued attribute (the impact of the attribute is relatively high), its levels (compared to the levels of other attributes) are not very far apart on the utility scale (cf. Table 7). On the other hand, the levels of "examination cost" are the furthest on the utility scale, but the attribute itself has a relatively weak impact (for this reason, the attribute "examination cost" is chosen as a reference). This illustrates the decisive advantage of the BWS method compared to the DCE method, which only estimates the differences between the attribute levels. Since BWS, DCE, CR measure the same construct - utility, it can be said that the convergent validity has partially confirmed the external validity of the BWS method (Leung, 2013). Pearson correlation coefficients show statistically significant ($p < 0.01$) very good positive correlation between the methods ($r > 0.7$).

4.5. Study limitations

The investigation at the county level (although the programme itself is conducted at the county level in coordination with the public health institute) is certainly a limitation. However, as this type of research requires the support of county public health institutes, a national level analysis would require a national level coordination that could not be provided in this case.

As mentioned earlier, the distribution of the sample by region is similar to the actual distribution of the population, but the educational structure of the sample is somewhat different from the actual educational structure in the Primorje-Gorski Kotar County. Namely, the percentage of respondents with completed secondary education (59.2%) is similar (according to the 2001 census) to the actual situation in the Primorje-Gorski Kotar County (54.1%). However, the educational structure of the sample (if we consider the levels of education lower and higher than high school) shows an advantage in favour of the respondents with higher levels of education (graduate and postgraduate education) (37.46%), which does not correspond to the actual situation. On the other hand, the percentage of population with primary education and less in the Primorje-Gorski Kotar County (Health Statistical Yearbook of the Primorje-Gorski Kotar County, 2011) was 30%, while the sample included only 3.34% of this group. Although this educational structure of the sample is a weakness in the research conducted, it is as it should be expected due to the process of filling out the questionnaire, namely the selection of the "best" and "worst" levels of characteristics, as well as the evaluation of the programmes offered based on their attributes. Completing the questionnaire might be a cognitively demanding task for those with low levels of education, which is the reason they might refuse to complete the questionnaire. This limitation could in future be overcome by the investigator's assistance (verbal instructions) to the respondents when completing the questionnaire.

Although, the sample size was sufficient for the preference elicitation, it could not provide a robust analysis of the relationship between the respondents' socio-economic characteristics and their screening preferences. Research limitation at the level of analysing homogeneous groupings based on socio-economic characteristics relates to an insufficiently large statistical sample for this type of extended model, suggesting implications for future research. Namely, future research should include a larger sample, which would allow for a robust analysis at the level of different groups, as well as an extension of the model with interactions of covariates and the attribute levels (in addition to only attribute interactions), which would provide an even more subtle insight into the preferences of different groups of respondents.

5. DISCUSSION AND CONCLUSION

In economics, the importance of preferences in consumer decision-making and of determining the value of goods is widely recognised. However, in health economics, health was the only relevant outcome in determining the benefits of health care services until the 1990s. Subsequently, the importance of non-medical and process attributes has been confirmed by several empirical studies (Trapero-Bertran et al., 2019; de Bekker-Grob et al., 2012; Protiere et al., 2004; Ryan, 1999; Donaldson & Shackley, 1997), according to which patients are willing to substitute health outcomes for various process attributes.

Decision making in health care (especially when it comes to public health policy and health care reform) is often based not only on the results of economic evaluations, but there are factors that equally require the attention of decision-makers. An OECD study (2005) has shown that the stakeholders' acceptance of health interventions, programmes, and policies is an important determinant of their practical success. More

recently, there has been a greater focus on public preferences in health decision-making in the European region. For example, Marsh et al. (2020) identified 22 European countries that use preference data (using different methods) for various purposes, ranging from assessing different impacts on patients to evaluating non-health factors on the payment system to estimating opportunity costs.

This paper confirms the theoretical justification and methodological applicability of the choice-based methods for preference evaluation. Their advantage is most evident in the selection process, which is similar to the real decision-making process. Indeed, most decisions in everyday life consist of choosing between comparable competitive options. Trade-offs that consumers make by choosing smaller quantities of one good for larger quantities of another reveal the central marginal value they attach to that good. Since this study was conducted, choice-based methods have been used in much health-related research to provide prevention and disease-specific decision-making information. From a prevention perspective, the BWS method has been used to inform primary prevention (healthy behaviour) issues, such as fresh fruit and vegetable consumption in Northwest Italy (Massaglia et al., 2019) or physical activity preferences in Australia (Franco et al., 2015) and Sweden (Aboagye et al., 2017); on issues of secondary prevention programmes, e.g. tuberculosis screening for healthcare workers in South Africa (O'Hara et al., 2015) or preferences for a genetic testing programme for colorectal cancer in Denmark (Veldwijk et al., 2016). In addition, BWS has been used in a multi-country (12 European and 8 non-European) study of key health outcomes for patients with spondyloarthritis (Kiltz et al., 2016), patient preference elicitation for different attributes of multiple sclerosis drug treatment in Canada (Lynd et al., 2016), and in a study of nutritional intervention in colorectal cancer survivors in England (Wright et al., 2017). In addition, a DCE study on preferences for Pap smear and mammography has been conducted in Singapore (Bilger et al., 2020). Two studies in Croatia on marginal willingness to pay estimated using BWS data for dental care (Sever, Verbič & Klarić Sever, 2020) and the national cervical screening programme (Dukić Samaržija, 2019) have been conducted in Croatia.

This study confirms that prevention programme attributes are significantly associated with respondents' utility (usefulness), which in turn is related to women's decision to participate in screening. Since the response of the target population is the crucial condition for achieving the expected benefits (in the sociological, epidemiological and economic sense), it is in direct function for a better adaptation and implementation of prevention programmes. In this sense, the research results clearly indicate the importance of considering the population's preferences in the design and implementation of prevention services in health care. As the Croatian cervical cancer screening programme requires a coordination of activities and interests at different levels (Ministry of Health, Public Health institutes, gynaecologists, and cytologists), public recognition of the preferences of the target population could be a good starting point for harmonising conflicting interests. This point has also been borne out in practice. After the implementation of the first cycle of the National Programme of Early Detection of Cervical Cancer, which began in December 2012, the programme was suspended (partly on the basis of our preference survey study) and is currently in the phase of reorganisation. In 2018, planning and preparations began for

the implementation of the redefined programme. It is expected that the protocols will improve the monitoring (by including them in the new payment scheme for primary care), which will allow an evaluation of the programme and health workers, and strengthen human resources and their capacity to successfully implement the programme (Croatian Institute of Public Health, 2019). The author plans to conduct a preference survey study after completion of the redefined national programme to evaluate its effectiveness in increasing the uptake of screening.

Although it is impossible to fully implement market settings in health care, it is still important to consider the preferences of the target population in order to achieve a greater response to prevention programmes, which is well recognised in the EU (de Lorenzo & Apostolidis, 2019). This paper provides valuable insights into the relative importance of national prevention programme attributes, i.e. respondents' willingness to substitute one attribute for another (attribute substitution rate). It thus provides an applicable contribution to public health decision-making. This does not mean that the decision to fund public programmes should be based solely on preference analysis, as it depends primarily on the epidemiology of the population, scientific advances in medical diagnostics, and the cost-effectiveness of the programme itself. However, since the response of the target population is an essential prerequisite for achieving the intended benefits, assessing the preferences of the target population is in direct function to better adapt and implement public health programmes.

REFERENCES

1. Aboagye, E., Hagberg, J., Axén, I., Kwak, L., & Lohela-Karlsson, M. (2017). Individual preferences for physical exercise as secondary prevention for non-specific low back pain: A discrete choice experiment. *PLoS ONE*, *12*(12). https://doi.org/10.1371/journal.pone.0187709
2. Adamowitcz, W., Louviere, J., & Swait, J. (1998). *Introduction to attribute-based stated choice methods* [Final Report,]. Alberta, Canada: NOAA Resource Valuation Branch Damage Assessment Centre. Retrieved from NOAA Resource Valuation Branch Damage Assessment Centre website: http://www.greateratlantic.fisheries.noaa.gov/hcd/statedchoicemethods.pdf
3. Bilger, M., Özdemir, S., & Finkelstein, E. A. (2020). Demand for Cancer Screening Services: Results From Randomized Controlled Discrete Choice Experiments. *Value in Health*, *23*(9), 1246–1255. https://doi.org/10.1016/j.jval.2020.06.004
4. Bockstael, N. E., & McConnell, K. E. (2007). *Environmental and resource valuation with revealed preferences, edt*. Dordrecht: Ian J. Bateman, Springer.
5. Bridges, J. (2003). Stated preference methods in health care evaluation: An emerging methodological paradigm in health economics. *Applied Health Economics and Health Policy*, *2*(4), 213–224.
6. Cerigo, H., Coutlée, F., Franco, L., E., & Brassard, P. (2013). Factors associated with cervical cancer screening uptake among Inuit women in Nunavik. *Public Health*, *13*(438). https://doi.org/10.1186/1471-2458-13-438

7. Clark, M. D., Determann, D., Petrou, S., Moro, D., & Bekker-Grob, E. W. (2014). Discrete choice experiments in health economics: A review of the literature. *Pharmacoeconomics*, *32*, 883–902. https://doi.org/10.1007/s40273-014-0170-x
8. Croatian Institute of Public Health (2019. (n.d.). *Epidemiology and Immunisation*. Retrieved from https://www.hzjz.hr/sluzba-epidemiologija-zarazne-bolesti/odrzana-prva-radionica-reorganiziranog-nacionalnog-programa-ranog-otkrivanja-raka-vrata-maternice/
9. De Bekker-Grob, E. W., Ryan, M., & Gerard, K. (2012). Discrete choice experiments in health economics: A review of the literature. *Health Economics*, *21*, 145–172.
10. De Lorenzo, F., & Apostolidis, K. (2019). The European Cancer Patient Coalition and its central role in connecting stakeholders to advance patient-centric solutions in the mission on cancer. *Molecular Oncology*, *13*, 653–666. https://doi.org/10.1002/1878-0261.12448
11. Donaldson, C., & Shackley, P. (1997). Does process utility exist? A case study of willingness to pay for laparoscopic cholecystectomy. *Social Science and Medicine*, *44*(5), 699–707.
12. Dukić Samaržija, N. (2019). Determining the Marginal Willingness to Pay for the Cervical Cancer Screening Program in Croatia: A Best-Worst Scaling Approach. *Society and Economy*, *41*(4), 433–447. https://doi.org/10.1556/204.2019.016.
13. Ekonomsko vrednovanje na strani ponude i potražnje—Ključ alokacije resursa u javnom zdravstvu. (2015). *Ekonomski pregled*, *66*(2), 177–200.
14. Espina, C., Soerjomataram, I., Forman, D., & Martín-Moreno, J. M. (2018). Cancer prevention policy in the EU: Best practices are now well recognised; no reason for countries to lag behind. *Journal of Cancer Policy*, *18*, 40–51. https://doi.org/10.1016/j.jcpo.2018.09.001
15. Finn, A., & Louviere, J. J. (1992). Determining the appropriate response to evidence of public concern: The case of food safety. *Journal of Public Policy & Marketing*, *11*(1), 12–25.
16. Flynn, T. (2010). Valuing citizen and patient preferences in health: Recent developments in three types of best–worst scaling. *Expert Review of Pharmacoeconomics & Outcomes Research*, *10*(3), 259–267.
17. Flynn, T., Louviere, J. J., Peters, T. J., & Coast, J. (2008). Estimating preference for a dermatology consultation using Best-Worst Scaling: Comparison of various methods of analysis. *Medical Research Methodology*, *8*(76). https://doi.org/10.1186/1471-2288-8-76
18. Fort, V. K., Makin, M. S., Siegler, A. J., Ault, K., & Rochat, R. (2011). Barriers to cervical cancer screening in Mulanje, Malawi: A qualitative study. *Patient Preference and Adherence*, *5*, 125–131.
19. Franco, M. R., K., H., C., S., H., F. P., J., R., L., G. J., & L, F. M. (2015). Eliciting older people's preferences for exercise programs: A best-worst scaling choice experiment. *Journal of Physiotherapy*, *61*(1), 34–41. https://doi.org/10.1016/j.jphys.2014.11.001
20. Guttman, R., Castle, R., & Fiebig, D. G. (2009). *Use of discrete choice experiments in health economics: An update of the literature*. Sydney: University of Technology. Retrieved

from http://www.chere.uts.edu.au/pdf/wp2009_2.pdf.

21. Hall, J., Viney, R., Haas, M., & Louviere, J. (2001). Using stated preference discrete choice modelling to evaluate health care programs. *Journal of Business Research*, *57*(9), 1026–1032.
22. *Health Statistics Yearbook of the County of Primorje-Gorski Kotar*. (2012). Retrieved from http://www.zzjzpgz.hr/statistika/statistika2012/index.html.
23. Holroyd, E., Twinn, S., & Adab, P. (2004). Socio-cultural influences on Chinese women's attendance for cervical screening. *Journal of Advanced Nursing*, *46*(1), 42–52.
24. Johnston, G., Orbell, S., Crombie, I., Robertson, A., & Kenicer, M. (1996). Women's preferences for cervical screening: Who, where and when. *Health Education Journal*, *55*(1), 84–91.
25. Kivistik, A., Lang, K., Baili, P., Anttila, A., & Veerus, P. (2011). Women's knowledge about cervical cancer risk factors, screening, and reasons for non-participation in cervical cancer screening programme in Estonia. *BMC Women's Health*, *11*(43). https://doi.org/10.1186/1472-6874-11-43
26. Kjær, T. (2005). *A review of the discrete choice experiment—With emphasis on its application in health care*. Syddansk Universitet, Health Economics Paper No.1.
27. Laviana, A. A., Luckenbaugh, A. N., & Resnick, M. J. (2020). Trends in the Cost of Cancer Care: Beyond Drugs. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *38*(4), 316–322. https://doi.org/10.1200/JCO.19.01963
28. Leung, T. T. W. (2013). *Validity and feasibility of best worst scaling using multiple treatment outcomes of Parkinson's disease* (Master thesis,). School of Health Science, University of Twente.
29. Li, S., Liu, S., Ratcliffe, J., Gray, A., & Chen, G. (2019). Preferences for cervical cancer screening service attributes in rural China: A discrete choice experiment. *Patient Preference and Adherence*, *13*, 881–889. https://doi.org/10.2147/PPA.S201913
30. Louviere, J. J., & Lancsar, E. (2009). Choice experiments in health: The good, the bad, the ugly and toward a brighter future. *Health Economics, Policy and Law*, *4*, 527–546.
31. Louviere, J. J., & Woodworth, G. G. (1983). Design and analysis of simulated consumer choice or allocation experiments: An approach based on aggregate data. *Journal of Marketing Research*, *20*, 350–367.
32. Louviere, J. L., Hensher, D. A., & Swait, J. D. (2000). *Stated choice methods: Analysis and applications*. Cambridge: Cambridge University Press.
33. Lynd, L. D., Traboulsee, A., Marra, C. A., Mittmann, N., & Evans, C. (2016). Quantitative analysis of multiple sclerosis 'patients' preferences for drug treatment: A best–worst scaling study. *Therapeutic Advances in Neurological Disorders*, 287–296. https://doi.org/10.1177/1756285616648060.
34. Mangham, L. J., & Hanson, K. (2009). How to do (or not to do)...Designing a discrete choice experiment for application in a low-income country. *Health Policy and Planning*, *24*, 151–158.
35. Marley, A. J., Flynn, T. N., & Louviere, J. J. (2008). Probabilistic models of

set-dependent and attribute level best-worst choices. *Journal of Mathematical Psychology*, *52*, 281–296.

36. Marsh, K., Til, J. A., Molsen-David, E., Juhnke, C., & N., H. (2020). Health Preference Research in Europe: A Review of Its Use in Marketing Authorization, Reimbursement, and Pricing Decisions-Report of the ISPOR Stated Preference Research Special Interest Group. *Value in Health*, *23*(7), 831–841. https://doi.org/10.1016/j.jval.2019.11.009.
37. Massaglia, S., Borra, D., Peano, C., Sottile, F., & Merlino, V. M. (2019). Consumer Preference Heterogeneity Evaluation in Fruit and Vegetable Purchasing Decisions Using the Best–Worst Approach. *Foods*, *8,266*. https://doi.org/10.3390/foods8070266
38. Matejic, B., Vukovic, D., Pekmezovic, T., Kesic, V., & Markovic, M. (2011). Determinants of preventive health behaviour in relation to cervical cancer screening among the female population of Belgrade. *Health Education Research*, *26*(2), 201–211.
39. McCaffery, K., & Irwig, L. (2005). Australian women's needs and preferences for information about human papillomavirus in cervical screening. *Journal of Medical Screening*, *12*(3), 134–141.
40. McFadden, D. (1973). Conditional logit analysis of qualitative choice behaviour. In P. Zarembka (Ed.), *Frontiers in Economics*. New York: Academic Press.
41. McIntosh, E. S. (2003). *Using discrete choice experiments to value the benefits of health care* ((Doctoral dissertation). Aberdeen,). University of Aberdeen.
42. Nicoll, P. M., Narayan, K. V., & Paterson, J. G. (1991). Cervical cancer screening: Women's knowledge, attitudes and preferences. *Health Bulletin*, *49*(3), 184–190.
43. Oberlin, A. M., Pasipamire, T., & Chibwesha, C. J. (2019). Exploring women's preferences for HPV-based cervical cancer screening in South Africa. *International Journal of Gynecology & Obstetrics*, *146*(2), 192–199.
44. O'Hara, N. N., Roy, L., O'Hara, L. M., Spiegel, J. M., & Lynd, L. D. (2015). Healthcare Worker Preferences for Active Tuberculosis Case Finding Programs in South Africa: A Best-Worst Scaling Choice Experiment. *PLoS ONE*, *10*(7), 0133304. https://doi.org/10.1371/journal.pone.0133304
45. Orme, B. K. (2006). *Getting started with conjoint analysis: Strategies for product design and pricing research*. Madison, Wisconsin: Research Publishers LLC.
46. Protiere, C., Donaldson, C., Moatti, J. P., Luchini, S., & Shackley, P. (2004). 'The impact of information on non-health attributes on willingness to pay for multiple health care programmes. *Social Science and Medicine*, *58*(7), 1257–1269.
47. Raginel, T., Grandazzi, G., Launoy, G., Trocmé, M., Christophe, V., Berchi, C., & Guittet, L. (2020). Social inequalities in cervical cancer screening: A discrete choice experiment among French general practitioners and gynaecologists. *BMC Health Services Research*, *20*(693). https://doi.org/10.1186/s12913-020-05479-w
48. Ryan, M. (1999). Using conjoint analysis to take account of patient preferences and go beyond health outcomes: An application to in vitro fertilisation.

Social Science & Medicine, *48*, 535–546.

49. Sassi, F., & Hurst, J. (2008). The Prevention of Lifestyle-Related Chronic Diseases: An Economic Framework. *OECD Health Working Paper*, (32).
50. Sever, I., Verbič, M., & Klaric Sever, E. (2020). Estimating Attribute-Specific Willingness-to-Pay Values from a Health Care Contingent Valuation Study: A Best-Worst Choice Approach. *Applied Health Economics and Health Policy*, *18*(1), 97–107. https://doi.org/10.1007/s40258-019-00522-2.
51. Society, A. C. (2018). *Economic Impact of Cancer*. Retrieved from https://www.cancer.org/cancer/cancer-basics/economic-impact-of-cancer.html
52. Soekhai, V., Bekker-Grob, E. W., & Ellis, A. R. (2019). Discrete Choice Experiments in Health Economics: Past, Present and Future. *Pharmacoeconomics*, *37*, 201–226. https://doi.org/10.1007/s40273-018-0734-2
53. Subramanian, S., Kaganova, Y., Zhang, Y., Hoover, S., Nyambe, N., Pinder, L., … Parham, G. (2018). Patient Preferences and Willingness to Pay for Cervical Cancer Prevention in Zambia: Protocol for a Multi-Cohort Discrete Choice Experiment. *JMIR Research Protocols*, *7*, 7. https://doi.org/10.2196/10429
54. Trapero-Bertran, M., Rodríguez-Martín, B., & López-Bastida, J. (2019). What attributes should be included in a discrete choice experiment related to health technologies? A systematic literature review. *PLoS ONE*, *14*, 7. https://doi.org/10.1371/journal.pone.0219905
55. Veldwijk, J., Lambooij, M., & Kallenberg, F. (2016). Preferences for genetic testing for colorectal cancer within a population-based screening program: A discrete choice experiment. *European Journal of Human Genetics*, *24*, 361–366. https://doi.org/10.1038/ejhg.2015.117
56. Wei-Hua, T., Chin-Shyan, C., & Tsai-Ching, L. (2010). The demand for preventive care services and it's relationship with inpatient services. *Health Policy*, *94*(2), 164–174.
57. Which aspects of health are most important for patients with spondyloarthritis? A Best Worst Scaling based on the ASAS Health Index. (2016). *Rheumatology*, *55*(10), 1771-1776,. https://doi.org/10.1093/rheumatology/kew238
58. Wordsworth, S., Ryan, M., Skåtun, D., & Waugh, N. (2006). Women's preferences for cervical cancer screening: A study using a discrete choice experiment. *International Journal of Technology Assessment in Health Care*, *22*(3), 344–350.
59. Wright, S. J., Gibson, D., & Eden, M. (2017). What are colorectal cancer survivors' preferences for dietary advice? A best-worst discrete choice experiment. *Journal of Cancer Survivorship*, *11*, 782–790. https://doi.org/10.1007/s11764-017-0615-2

KORIŠTENJE METODA ISKAZANIH PREFERENCIJA ZA INFORMIRANJE U ODLUČIVANJU O JAVNOM ZDRAVSTVU: DOKAZI IZ HRVATSKE

Sažetak

Hrvatska je prepoznala značaj prevencijskih programa u području javnog zdravstva, iako njihova učinkovitost nije zadovoljavajuća zbog niskog odziva ciljne populacije, a što ima negativan učinak na racionalizaciju javne potrošnje. Jedno od mogućih rješenja je razmatranje iskazanih preferencija ciljne populacije. Stoga je temeljni cilj ove studije evaluirati korištenje metode iskazanih preferencija u unapređenju preventivnih programa javnog zdravstva. U svrhu provedbe studije, izrađen je upitnik, korištenjem triju različitih metoda – eksperimenta diskretnog odabira, uvjetnog vrednovanja te metode najboljeg i najgoreg odabira. Studija pokazuje da su atributi hrvatskog programa za rano otkrivanje raka maternice značajno povezani s razinom korisnosti (zadovoljstva) korisnika programa, a što je povezano s odgovorom žena na anketni upitnik. S obzirom da sve tri korištene metode mjere isti konstrukt – korisnost, može se reći da konvergentna validnost djelomično potvrđuje eksternu validnost navedenih metoda. Autorica zaključuje da je u procesu planiranja, implementacije i re-evaluacije programa javnog zdravstva nužno primijeniti tržišna načela, kroz analizu potražnje (ciljne populacije), korištenjem metoda iskazanih preferencija.

Ključne riječi: *javno zdravstvo, metode zasnovane na izboru, prevencija, odlučivanje, skaliranje najboljeg i najgoreg odabira*