Development of Willingness-to-Pay Questionnaire on Health Insurance Cost-Sharing for Catastrophic Prescription Drugs: A Content Validity and Reliability Assessment

Diesty Anita Nugraheni[1,2], Satibi[3*], Susi Ari Kristina[3] and Diah Ayu Puspandari[4]

[1.] Doctoral Graduate Program, Faculty of Pharmacy, Universitas Gadjah Mada, Yogyakarta, Indonesia, 55281
[2.] Department of Pharmacy, Faculty of Mathematics and Natural Sciences, Universitas Islam Indonesia, Yogyakarta, Indonesia, 55584
[3.] Department of Pharmaceutics, Faculty of Pharmacy, Universitas Gadjah Mada, Yogyakarta, Indonesia, 55281
[4.] Department of Health Policy and Management, Faculty of Medicine, Public Health, and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia, 55281

Article Info	ABSTRACT
	Prescription drug cost-sharing in Indonesia remains challenging because no valid willingness-to-pay (WTP) questionnaire on cost-sharing is available. This study aimed to conduct an item selection and assess the content validity of a WTP questionnaire for prescription drug cost-sharing under the national health insurance scheme in patients suffering from catastrophic illnesses. The method was a cross-sectional study and used a four-step design, including study development, content validity, response process, and reliability. This research involved nine experts, including professionals, academics, and pharmacy experts in hospitals, to validate the content of a WTP questionnaire for prescription drug cost-sharing. The questionnaire items were extracted from many sources, such as cost analysis of catastrophic prescription drug data in Indonesia and reviews of articles with the terms "willingness-to-pay," "cost-sharing," "health insurance," "prescription," and "questionnaire." This content validity study used the content validity index (CVI) and content validity ratio (CVR) for both the individual item measurement and the overall scale. The Pearson correlation was assessed for the test–retest reliability. A draft questionnaire of 47 items was developed with four domains: healthcare utilization information, participation in health insurance information, drug information, and cost-sharing scenarios. Three items were removed, and three were merged based on the content validity study. The overall scale of the questionnaire was good with 0.94 S-CVI/Ave. The test–retest reliability coefficient had a strong correlation (r = 0.98). We conclude that the WTP questionnaire for prescription drug cost-sharing had been developed and validated through evidence validity, including content validity (CVI and CVR), response process, and reliability. As a follow-up, a new 42-item questionnaire was developed. **Keywords:** Willingness-to-Pay, Content Validity, Prescription Drug Cost-Sharing, Health Insurance, Questionnaire

INTRODUCTION

Indonesia is a developing country in Asia. The government launched the Indonesian version of universal health coverage in 2014, titled Jaminan Kesehatan Nasional (JKN) or the National Health Insurance (NHI) program, to protect all citizens (Agustina *et al*., 2019; World Health Organization, 2010). However, after four years of implementation by the Healthcare and Social Security Agency (BPJS-Kes), the JKN scheme faced a budget deficit due to various factors, including JKN members did not pay the premiums, the JKN premiums failed to meet the actuarial values, the JKN scheme did not implement cost-sharing but

offered unlimited benefit packages, and the costs of catastrophic illnesses were quite high.

It is important to address such financing problems with quality and cost control efforts, one of which is cost-sharing (Koohi Rostamkalaee *et al*., 2022) (Seaman *et al*., 2017) (Park & Choi, 2020). Cost-sharing is a technique to control healthcare costs in which patients must pay a portion of the healthcare services (Remler & Greene, 2009). Some studies in developed countries showed that healthcare utilization significantly increased as a result of implementing cost-sharing (Hossein & Gerard, 2013; Salampessy *et al*., 2018). Unfortunately, there have not been many studies on the effect and implementation of cost-sharing in developing countries (Feng *et al*., 2020; Nguyen & Connelly, 2017). It is important for health insurance providers to decide on the premiums that members should pay to guarantee their financial sustainability, especially in low- and middle-income countries, in which the health insurance coverage should range between 20 to 60% and the insurance members are still required to spend out-of-pocket (OOP) expenses (World Health Organization, 2010). Therefore, collecting valid information on the amount of cost-sharing patients are willing to pay for their health insurance programs is crucial (Nosratnejad *et al*., 2014). However, there have been no studies on the precise amount people are willing to pay for the cost-sharing of social health insurance.

Willingness-to-pay (WTP) is the maximum amount of payment that a person is willing to spend for a product or service depending on the perceived utility and values (Breidert, 2005). Various factors affect the WTP for health insurance, including health factors, sociodemographic characteristics, and insurance-related knowledge (Gidey *et al*., 2019; Minyihun *et al*., 2019), as well as environmental and economic factors (Minyihun *et al*., 2019). WTP can be measured through a direct interview, phone call, mail, or survey using a questionnaire (Rascati, 2014). WTP measurement can be performed using a questionnaire that offers a certain amount of money along with a higher or lower amount, which usually depends on the responses to the first amount offered (Adebayo *et al*., 2015; Darmawan *et al*., 2019; Gidey *et al*., 2019; Thi Thuy Nga *et al*., 2018).

This paper examined the validity and reliability of a WTP questionnaire to establish prices for prescription drug cost-sharing of catastrophic illnesses in Indonesia. We suggest that a WTP questionnaire may become a useful technique for the payer (BPJS Kesehatan) to estimate demand for services at the given target prices. Prescription drug cost-sharing in Indonesia remains challenging because no valid questionnaire on WTP is available. The questionnaire can be helpful for health-related policymakers seeking data on WTP for cost-sharing. We found two other studies of predictive validity and pilot study of WTP questionnaire technique in a developing-country setting (Asenso-Okyere *et al*., 1997; Foreit & Foreit, 2003). Foreit *et al*. found that the WTP questionnaire technique was reliable and valid for predicting WTP values when used for reproductive service prices in developing countries (Foreit & Foreit, 2003). To the best of our knowledge, a validated questionnaire on WTP for prescription drug cost-sharing under NHI has yet to be developed. This paper assessed the validity and reliability of a WTP questionnaire on prescription drug cost-sharing under NHI in developing countries. This study of WTP aimed to help governments set or increase prices charged for health insurance payments for prescriptions for catastrophic illnesses.

Five general sources of evidence support the validity of a research tool: 1) content, in which the items of a research tool completely represent the measured outcomes, 2) response process, where subjects correctly understand the items of a research tool, 3) internal structure, in which tolerable reliability and factor structure are shown, 4) relations to other variables, where there is a correlation among the tools that measure similar outcomes, and 5) consequences, in which interpretation of scores can justify a difference (Cook & Beckman, 2006; Messick, 1995). Evidence should be pursued from different sources to support any interpretation, and strong evidence from one source does not negate the need to find evidence from other sources. Reliable scores are necessary but insufficient for valid interpretation (Cook & Beckman, 2006). In our study, we examined three sources of evidence of a new instrument: content validity, response process validity, and internal structure for reliability test.

Developing a valid, reliable measurement instrument is a highly critical point. A validity test needs to be performed before a reliability test to ensure that what is tested is a valid instrument. Validity guarantees that the instrument measures what it anticipates, determines, and reflects the intended theoretical concept. Evaluation of the content validity of an instrument is the vital and critical early stage in determining the validity of

instruments (Polit *et al.*, 2007). Meanwhile, content validity refers to the level at which the items of an instrument sufficiently represent the domains of the content. If an instrument has inadequate content validity, it is impossible to develop its reliability (Moshki *et al.*, 2020; Zamanzadeh *et al.*, 2014).

The study aimed to conduct validity and reliability tests of the WTP questionnaire for prescription drug cost-sharing under the NHI scheme for patients suffering from catastrophic illnesses who received either outpatient or inpatient services with their healthcare costs covered by the NHI. Studies focusing on survey instrument validation for the measurement of WTP for prescription drug cost-sharing among Indonesian citizens who suffer from catastrophic illnesses have not been widely available. In addition, this study specifically aimed at identifying the content validity index (CVI) and the content validity ratio (CVR) of each item in the questionnaire and calculating the questionnaire's mean scale content validity index (S-CVI/Ave). The CVI was chosen for its comprehensibility and ease of communication and calculation (Polit *et al.*, 2007), and it was measured for the proportion of the items that are rated relevant by the panelists (Farzanegan *et al.*, 2017). In addition, CVR indicates whether an item should be included or not in the questionnaire. As shown in some previous studies, the content validity of an instrument can improve data effectiveness and minimize measurement errors (Polit & Beck, 2006; Sangoseni *et al.*, 2013). It also indicates the extent to which a certain instrument has suitable items representing the measured constructs (Polit & Beck, 2006). Many previous studies have used CVI (Baker *et al.*, 2017; Chiwaridzo *et al.*, 2017; Farzanegan *et al.*, 2017; Juengst *et al.*, 2019; Lavoie Smith *et al.*, 2017; Moghaddam *et al.*, 2019; Moshki *et al.*, 2020; Saber *et al.*, 2017) to increase the validity of their measurement instrument because it indicate the extent to which the selected items contain appropriate samples for the contents that should be measured using the questionnaire (Polit *et al.*, 2007; Polit & Beck, 2006).

MATERIALS AND METHODS

A quantitative study was conducted in July and August 2021 in Yogyakarta, Indonesia. This research was intended to conduct a content validation of the developed questionnaire. In general, the study used a four-step design: 1) development of questionnaire, 2) content validity study including CVI and CVR, 3) response process, and 4) reliability assessment including test–retest reliability. This study received approval from the Ethics Committee of the Faculty of Medical Sciences, Universitas Gadjah Mada, Yogyakarta, Indonesia (approval code: KE/FK/0539/EC/2021).

Eligibility criteria for the validity study

The content validity of an instrument can be defined using the viewpoints of a panel of experts who are familiar with the conceptual baseline of the instrument (Grant & Davis, 1997); thus, they are referred to as content experts (Farzanegan *et al.*, 2017). Content experts have worked in the field or published related articles. In this study, the criteria for selecting these experts were their number of publications on WTP or work experience in a hospital pharmacy. In addition, selecting experts in measurement or related fields can be beneficial for determining whether the new measurement is well-developed and appropriate for psychometric testing (Davis, 1992). The exclusion criteria in this research were those having less than five years of experience, not being involved in services of JKN members, or having no expertise related to JKN financing. A total of ten experts were asked to participate in this research, but one canceled her participation before filling out the content validity form. Two professionals are from academia with doctoral degrees, have researched WTP, and have NHI publications. The other seven experts are from nonacademic fields of work in hospital pharmacies.

Step 1: Development of Questionnaire

The literature review was done using ScienceDirect, PubMed, Google Scholar, and Scopus without any date of publication limitations with several keywords, including "willingness-to-pay," "health insurance," "cost-sharing," "prescription," and "questionnaire." Duplicate items and previous studies of health insurance systems not relevant to this study (due to specific contexts unable to be generalized to the context in the current study) were excluded. In developing the initial version of the questionnaire, the authors collected all the factors of WTP studied in some articles (Adams *et al.*, 2015; Akwaowo *et al.*, 2021; Gidey *et al.*, 2019; Haile *et al.*, 2014; Minyihun *et al.*, 2019; Nosratnejad *et al.*, 2014). In addition, the questionnaire was developed by using observational data on prescription costs to establish the bid amount of WTP. Finally, the questionnaire comprised 57 items categorized into five domains, including sociodemographic characteristics.

Questionnaire on WTP for health insurance cost-sharing

The questionnaire on WTP was in Indonesian and translated into English for this manuscript (Nugraheni *et al*., 2022). The surveys used similar elicitation methods (a series of three closed-ended questions to cover the range of target prices, followed by a single open-ended question to elicit maximum WTP) and additional information on healthcare and health insurance as a factor related to participants' selection of WTP values. Parts I–IV contained questions to determine the factors that affect WTP. Part I explored the respondents' sociodemographic factors (i.e., sex, marital status, age, educational level, family size, and religion). Part II asked about healthcare information (chronic or acute illnesses that the respondents or their family members suffer from, place of care of the latest episode, more recent healthcare costs, and satisfaction with the quality and financing of the services provided by the more recent place of care). Part III asked about health insurance information (health insurance knowledge, membership, and premiums for the latest health insurance). These factors in Parts I, II, and III were included in the questionnaire because, as shown in previous studies, they were found to have a significant effect on WTP for health insurance (Adebayo *et al*., 2015; Chanie & Ewunetie, 2020; Darmawan *et al*., 2019; Deksisa *et al*., 2020; Gidey *et al*., 2019; Kado *et al*., 2020; Mekonne *et al*., 2020; Minyihun *et al*., 2019). Part IV asked about information on drug use (prescription drugs, drugs received by the patients, and drugs bought by the patients, both prescription and over-the-counter drugs, apart from those covered by the health insurance). A drug information section was added to the questionnaire to evaluate whether patients received the prescription drugs and the OOP spending for drugs. Part V presented three hypothetical cost-sharing scenarios that were modified from previous studies (Akwaowo *et al*., 2021; Gidey *et al*., 2019; Haile *et al*., 2014): scenario A for no insurance (out-of-pocket model), scenario B for social health insurance (present model), and scenario C for health insurance with prescription drug cost-sharing. In addition to patients' WTP for the scenario, Part V also provided cost-sharing method alternatives, namely deductible, copayment, or coinsurance, and offered a bid amount to determine the cost-sharing amount under the NHI. The bid amount of WTP was determined by conducting an observational study to find out the average cost of prescribed catastrophic drugs in Indonesia.

Step 2: Content Validation

After developing the first version, we consulted with a panel of experts on the questionnaire. In detail, content validation should go through several stages: preparation of a content validation form, selection of a panel of experts to review the questionnaire, review of both items and their domains, scoring each item, and CVI calculation (Yusoff, 2019). For the CVI calculation, the rating of an item should consider simplicity, relevance, and transparency using a four-point scale to avoid any neutral or unspecified responses (Ayre & Scally, 2014; Polit *et al*., 2007), namely 1 = not relevant, 2 = somewhat relevant, 3 = quite relevant, and 4 = highly relevant (Davis, 1992; Polit & Beck, 2006). The questionnaire was sent via email to the panel of experts. These experts were required to fill out a self-administered questionnaire and to return it within one week. Once finished, they were asked to provide written comments that were used to revise the items to improve their relevance to the targeted domains and checked its compliance with the writing rules. Prior to calculating the CVI, the relevance of an item had to be rated either as 1 (when the relevance is either 3 or 4) or as 0 (when the relevance is either 1 or 2). To determine the item-level CVI (I-CVI) in this study, the total CVI from all the experts was calculated then divided by the total number of experts. S-CVI/Ave is the average of the I-CVIs of an instrument or the average of item relevance based on the rating given by all of the experts. An I-CVI > 0.79 means that the item is appropriate, an I-CVI between 0.7 and 0.79 means that modification is needed, and an I-CVI < 0.7 means that the item should be removed (Lawshe, 1975; Moghaddam *et al*., 2019; Zamanzadeh *et al*., 2014). When four experts or more are involved in rating the validity of each item, the I-CVI should be 1.00 (Mohamad Marzuki *et al*., 2018; Yusoff, 2019). In this study, the items with an I-CVI ± 0.79 were revised and those with I-CVI < 0.70 were removed. Meanwhile, a scale with S-CVI/Ave higher than 0.9 is considered to have good content validity (Polit *et al*., 2007). The remaining items were then subjected to CVR, which is also useful for selecting the most relevant questions (Lawshe, 1975; Moghaddam *et al*., 2019). In this study, the nine experts were asked to determine if an item was relevant to measure a construct by specifying each of the questionnaire items as "essential, useful but inessential, or

unnecessary." A questionnaire was provided to an expert panel who was asked to score each of the items, ranging from 0 (unnecessary), 1 (unnecessary but helpful), to 2 (necessary). The calculation of the CVR used the following formula: CVR = (Ne-N/2)/(N/2), where Ne is the total number of experts saying that an item is necessary, and N is the total members of experts in a panel. If the average CVR is ≥0.62, then the questionnaire item is accepted with an acceptable significance (Lawshe, 1975; Zamanzadeh *et al.*, 2014).

Step 3: Response Process Validation

Other terms for response process validity are surface validity, appearance validity, or logical validity (Lau *et al.*, 2018). Response process validity is defined as "Do respondents understand the questions to mean what we intend them to mean?" Research showed that 14% of validation studies are based on participants' response processes. The response process validation of this study was assessed by interviewing ten participants selected from the targeted population. Interviews are a researcher's favorite validation method based on a response process (Padilla & Benítez, 2014). These participants evaluated the instructions, the response format, and the questionnaire items for clarity. In the response process, we recorded participants' difficulties and identified potential themes.

A thematic analysis was performed of the participants' responses. Two authors read all the answers and categorized the contents to determine how often different themes were found. Theme analysis of participants' information classifies emerging themes and concepts that allow authors to propose and change the administered measures (Padilla & Benítez, 2014).

Step 4: Reliability

The test–retest reliability was used to assess temporal reliability. Twenty participants (67% response rate) answered the questionnaire over a month interval. The participants with the same criteria as the research subjects were asked to be involved in the study that required them to complete a questionnaire. They took the questionnaire twice during their routine check-ups every month. Previous research took a relatively long inter-investigation time (weeks to months) (Bourke *et al.*, 2004; Lee *et al.*, 2020; Rozgonyi *et al.*, 2021). The interval period of two to four weeks is the longest recommended time interval for test–retest reliability (Kawakami *et al.*, 2020; Rozgonyi *et al.*, 2021).

The results of the two tests were correlated to determine the consistency using Pearson's correlation (Vaz *et al.*, 2013). The correlation between two tests indicated the instrument stability. The higher (closer to 1) level of correlation in the first and second tests, the better the reliability of the measurement scale. The Pearson correlation coefficient considered a strong, moderate, and poor correlation, respectively, are >0.6, 0.4–0.6, and <0.4 (Paiva *et al.*, 2014).

RESULTS AND DISCUSSION

The first version of the questionnaire had 47 items categorized into four parts (the sociodemographic part was excluded). This version was then sent to the experts for review. After this analysis, the suggestions were followed up on, resulting in the second version of the survey instrument, which consisted of 42 items. In the second analysis, 100% of the questions were rated relevant, and there were no suggestions for adding or revising items. This step resulted in the instrument's third version, which had 42 items. Nine experts were involved in this study with the following demographic characteristics (Table I).

The experts' average age and length of work experience were 39.9 (SD = 5.1) and 13.3 (SD = 7.2) years, respectively. Five experts have master's degrees, and the remaining two have doctorates. Seven experts work in a large hospital and have an office position as the head of the Pharmacy Department and Hospital Management. All were female experts, which is to be expected given that most pharmacists in Indonesia are female. One expert was a university lecturer and worked in hospital management. The remaining expert was a university lecturer with expertise is in WTP and health financing.

The content validation obtained through the experts' self-administered questionnaire (Table II). The distribution of the experts' assessment of 47 items (Table II). Each expert has different CVIs ranging from 0.79–1.00. Only three experts answered unrelated or slightly related questions. Most rated then either three or four, meaning most of the items were categorized as highly relevant and quite relevant questions.

All the experts were requested to rate the item relevance in the questionnaire, namely whether or not the questionnaire items were relevant to the object of measurement and to achieving the objectives of the questionnaire.

Table I. Demographic Characteristics of the Experts

Participant	Age (years)	Gender	Field of expertise	Education level	Position	Years of work experience	Affiliation
1	36	Female	Pharmacy	Master's Degree	Head of Pharmacy Department	12	Public Hospital
2	42	Female	Pharmacy	Doctorate Degree	Head of Faculty Department	11	University
3	40	Female	Pharmacy	Bachelor's Degree	Head of Pharmacy Department	14	Private Hospital
4	36	Female	Pharmacy	Master's Degree	Head of Pharmacy Department	12	Public Hospital
5	36	Female	Pharmacy	Master's Degree	Head of Pharmacy Department	6	Public Hospital
6	36	Female	Pharmacy	Bachelor's Degree	Head of Pharmacy Department	12	Public Hospital
7	43	Female	Pharmacy	Master's Degree	Head of Pharmacy Department	22	Public Hospital
8	40	Female	Pharmacy	Master's Degree	Hospital Management	5	Private Hospital
9	51	Female	Pharmacy	Doctorate Degree	Hospital Management	27	Private Hospital and University

UGM: Universitas Gadjah Mada

Table II. Content validity results of the Questionnaire on Willingness to Pay for cost-sharing by the experts.

Expert number	Total number of highly or quite relevant questions (n)	Total number of irrelevant or slightly relevant questions (n)	Total number of unanswered questions (n)	Total number of questions (n)	Total number of highly or quite relevant questions to the total number of questions (CVI)
1	44	3	0	47	0.94
2	47	0	0	47	1.00
3	37	10	0	47	0.79
4	47	0	0	47	1.00
5	47	0	0	47	1.00
6	40	7	0	47	0.85
7	47	0	0	47	1.00
8	47	0	0	47	1.00
9	42	5	0	47	0.89

*CVI=Content validity index

The experts were also asked to provide written recommendations (regarding removal or revision) for each item with a low rating (1 or 2). For the items that should be revised, they were asked to provide written comments on the clarity (the extent to which the item was structured clearly) and a possible alternative to ensure item relevance. These experts were expected to return the questionnaire review within one week at the most

through email. Fortunately, in this study, all the experts returned the forms, and only two of nine respondents took more than a month to return the form.

The I-CVIs for the questionnaire items in this study were 0.67–1.00. The I-CVIs of 27 out of 47 questions in the questionnaire were 1.00, indicating total agreement among the experts (Table III). Meanwhile, the questionnaire had S-

CVI/Ave of 0.94. Three items with I-CVI > 0.89 were merged or retained based on recommendations from the experts. Three items with I-CVI > 0.79 were revised based on the experts' suggestions. Three items were removed because of a disagreement among the experts with CVR < 0.62. The removed items (Table III) were as follows: two items in the healthcare information domain, consisting of information on acute illnesses that the respondents and their family members suffer from, and one item under the health insurance information domain, related to the source of information, which is the source where the respondents obtain JKN information from. These items were unnecessary because the research had clearly defined the inclusion criteria, one of which is patients suffering from catastrophic illnesses including heart diseases, cancer, and stroke. In addition, the sentence structures of items 11, 19, and 26 were rearranged and changed based on the experts' recommendations.

Based on the experts' review, questions with a CVR ≥ 0.62 could be retained, while those with a CVR < 0.62 could not. Therefore, the items were developed in a way that they met the expected CVR (≥0.62) (Saber *et al.*, 2017). The results showed CVR ranging from 0.3–1 and CVI ranging from 0.67–1. Therefore, not all items were valid. Items 2, 3, and 16 obtained a CVR lower than the standards. These items were then removed from the questionnaire because they were considered irrelevant and unimportant for measuring WTP for JKN cost-sharing among patients with catastrophic illnesses.

CVI is a technique that is most frequently used in research (Polit & Beck, 2006). Lynn (1986) recommended involving a minimum of three experts to a maximum of ten experts. This study involved nine experts; eight are hospital pharmacists and one is a senior lecturer (associate professor). The CVI in this study was calculated based on the ratings given by the experts; the range of 0.8–1 would be considered highly relevant (Lynn, 1986). I-CVI has to be 1.00 if five or fewer experts are involved, while if six to eight experts are involved, the I-CVI should be >0.83 (Polit *et al.*, 2007; Polit & Beck, 2006).

The Scale-CVI/Average for the 47 items was 94% (>90%), meaning that the items had good content validity (Davis, 1992; Polit *et al.*, 2007). However, Fallahzadeh *et al.* (2015) proposed a mean content validity of 90% when three experts are involved (Fallahzadeh *et al.*, 2015). In other words, the questionnaire contained relevant questions to measure WTP for prescription drug cost-sharing under JKN and to identify the factors affecting the WTP. These results were obtained after being reviewed by nine experts who represent the professions from academia and pharmacists in hospital although there were only two professionals who participated in this study. Not many professionals participated in this study because of their tight schedules at the time this study was carried out, i.e., August 2021, when the Covid-19 pandemic hit. To anticipate this, the experts were given one week to return the completed questionnaire, and several reminders were sent through emails. For any future studies with similar topics and respondents (experts), it is better that the experts are provided with adequate time (>1 week) to respond to the invitation to participate in a study.

CVI can show information at an item level regarding the extent to which there is an interrater agreement on the item relevance to achieve the objectives of the questionnaire. In this study, this information was used as the basis to revise or remove items. According to the experts, the questionnaire in general was already relevant to the study's objective. However, the sentence structures of some items had to be modified based on the experts' recommendations. Item 22, referring to item 24 "If No, how many drugs did you buy using your own money (not covered by BPJS) without any prescriptions in the last six months?" Since there was no correlation with item 24, i.e., "Did you g*et al*l the prescription drugs in the last six months?," it was revised to "How many drugs did you buy using your own money without any prescription in the last six months?" to meet the context of the research. Similar questions under a different domain were revised to make it clearer and meet the research context. Questions "How did you pay for your healthcare costs?" and "What kinds of health insurance that you had covered your health spending?" had similar answer choices with different contexts. These questions were then revised to "Who paid for your healthcare costs?" and "in addition to BPJS Kesehatan, what kinds of health insurance that you had to cover your health spending?" Meanwhile, items four to six were merged into one question, i.e., "How long had you been suffering from the last illness?"

Table III. Content Validity Results of the Questionnaire on Willingness-to-Pay for Cost-Sharing Based on the Questions (Excluding Questions on Demographic Information)

Domains	Questions	The number of experts who agreed that the questions were highly or quite relevant (n)	The number of experts who agreed that the questions were irrelevant or slightly relevant (n)	The number of experts who did not answer the questions (n)	CVR	I-CVI	Expert decision
Healthcare information	1	9	0	0	1.0	1.0	
	2	6	3	0	0.3	0.67	Removed
	3	7	2	0	0.6	0.78	Removed
	4	8	1	0	0.8	0.89	Merged
	5	8	1	0	0.8	0.89	Merged
	6	8	1	0	0.8	0.89	Merged
	7	8	1	0	0.8	0.89	
	8	9	0	0	1.0	1.0	
	9	8	1	0	0.8	0.89	
	10	9	0	0	1.0	1.0	
	11	8	1	0	0.8	0.89	Revised
	12	9	0	0	1.0	1.0	
	13	8	1	0	0.8	0.89	
	14	9	0	0	1.0	1.0	
	15	8	0	0	1.0	1.0	
Health insurance information	16	7	2	0	0.6	0.78	Removed
	17	9	0	0	1.0	1.0	
	18	9	0	0	1.0	1.0	
	19	9	0	0	1.0	1.0	Revised
	20	9	0	0	1.0	1.0	
	21	9	0	0	1.0	1.0	
Information on drug use	22	9	1	0	0.8	0.89	
	23	9	0	0	1.0	1.0	
	24	8	1	0	0.8	0.89	
	25	9	0	0	1.0	1.0	
	26	9	0	0	1.0	1.0	Revised
WTP for cost-sharing scenario	27	8	1	0	0.8	0.89	
	28	9	0	0	1.0	1.0	
	29	8	1	0	0.8	0.89	
	30	9	0	0	1.0	1.0	
	31	9	0	0	1.0	1.0	
	32	9	0	0	1.0	1.0	
	33	8	1	0	0.8	0.89	
	34	8	1	0	0.8	0.89	
	35	8	1	0	0.8	0.89	
	36	8	1	0	0.8	0.89	
	37	8	1	0	0.8	0.89	
	38	8	1	0	0.8	0.89	
	39	9	0	0	1.0	1.0	
	40	9	0	0	1.0	1.0	
	41	9	0	0	1.0	1.0	
	42	9	0	0	1.0	1.0	
	43	9	0	0	1.0	1.0	
	44	9	0	0	1.0	1.0	
	45	8	1	0	0.8	0.89	
	46	9	0	0	1.0	1.0	
	47	9	0	0	1.0	1.0	
			S-CVI/Ave			0.94	

**WTP= Willingness to pay; CVR= Content validity ratio; I-CVI=Item level-Content validity index

Table IV. Modified Questionnaire Items after the Response Process

The question	Difficulty encountered	The modified question
The respondent or his/her family members' experiences of suffering from a chronic illness	Multiple interpretations about chronic illness	The respondent or his/her family members' experiences of suffering from a chronic illness (for example: stroke, heart disease, cancer, asthma, diabetes mellitus, hypertension, kidney failure)?
How to pay for healthcare costs	Multiple interpretations about healthcare costs: whether it is only the costs not covered by the insurance or all the healthcare costs incurred (both covered and not covered by the insurance)	How to pay for the current healthcare costs (both covered and not covered by the insurance)
How to deal with financial problems to pay for healthcare costs	Multiple interpretations about how to deal with financial problems: whether it is difficulty to pay the insurance premiums or difficulty to pay both the insurance premiums and healthcare costs	How to deal with the financial problems to pay for healthcare costs, including to pay the insurance premiums and other healthcare costs
Level of satisfaction with the quality of healthcare service	Difficulty in understanding the definition of quality of healthcare service	Level of satisfaction with the quality of healthcare service (for example: performance speed, friendly staff, ease of administration, and services given by staff)
Level of satisfaction with healthcare costs	Multiple interpretations about costs	Level of satisfaction with healthcare costs, including the payment for the insurance premium and other healthcare costs

The next stage of validation evidence is the response process. It showed that the respondents were confused by the questionnaire (Table IV). Therefore, modifications were performed, resulting in the final version of the questionnaire on WTP for health insurance cost-sharing. We conducted an interview to directly obtain response process validity from the respondents. The participants were allowed to express their feelings, thoughts, opinions, etc.

The last stage of validation evidence is the test–retest reliability. The first (test) and second (retest) scores of the 20 participants were used for calculation. We used the Pearson correlation, which is commonly used to indicate relative reliability (Vaz *et al.*, 2013). The value of the Pearson correlation (r) obtained was 0.98 (correlation is strong if the value is from 0.6 to 1). The Pearson correlation coefficient was used to evaluate the correlations between two variables (Vartanian *et al.*, 2006). After testing the validity and reliability, a final questionnaire was obtained comprising 42 items divided into four parts,

namely, healthcare information (11 items), health insurance information (5 items), information on drug use (5 items), and WTP for cost-sharing scenarios (21 items).

Limitations

This study only involved experts in hospital pharmacy and two professionals from academia. It would have been better if more experts from more fields in the hospital had been involved. However, this was not possible due to the time limitations. It is necessary to interpret the results of this study carefully in light of its limitations. This is because the current study only involved pharmacy experts, so the results might not represent experts' opinions in other fields. Regarding these limitations, conducting another content validation study of the survey questionnaire with the same context is recommended by involving more content experts from various professions; each of them is asked to assess the fundamental aspects of the questions, including relevance, simplicity, ambiguity, and clarity. A larger size of samples is required to test

the construct validity, including the factor structure which was not carried out in this study. Finally, other types of construct validity may have to be performed, including relations to other variables and consequences.

CONCLUSION

The questionnaire developed in this study was valid and reliable for measuring WTP for prescription drug cost-sharing under NHI among patients with catastrophic illnesses. The questionnaire can be used in WTP studies in the same context but in different countries by adjusting the bid amount of WTP for each country. It is recommended for future studies to modify the questionnaire on other diseases by using a bid amount according to the prescribing cost. In the future, this WTP survey research can be used by the government and BPJS Health to determine JKN financing policies.

The study found some support for the construct validity: content validity (CVI and CVR), response processes, and test–retest reliability of the WTP questionnaire for prescription drug cost-sharing among JKN patients with catastrophic illnesses. This research indicated acceptable levels of validity, thereby illustrating a valid and reliable tool to be used in a survey of WTP for prescription drug cost-sharing. Further research is needed to clarify the construct validity of this instrument's other measures, such as relations to other variables and consequences.

ACKNOWLEDGMENT

AUTHORS' CONTRIBUTIONS

All authors had equally contributed with this study work.

CONFLICTS OF INTEREST

All authors have none to declare.

REFERENCES

Adams, R., Chou, Y.-J., & Pu, C. (2015). Willingness to participate and Pay for a proposed national health insurance in St. Vincent and the grenadines: A cross-sectional contingent valuation approach. BMC Health Services Research, 15, 148. https://doi.org/10.1186/s12913-015-0806-3

Adebayo, E. F., Uthman, O. A., Wiysonge, C. S., Stern, E. A., Lamont, K. T., & Ataguba, J. E. (2015). A systematic review of factors that affect uptake of community-based health insurance in low-income and middle-income countries. BMC Health Services Research, 15(1). https://doi.org/10.1186/s12913-015-1179-3

Agustina, R., Dartanto, T., Sitompul, R., Susiloretni, K. A., Suparmi, Achadi, E. L., Taher, A., Wirawan, F., Sungkar, S., Sudarmono, P., Shankar, A. H., Thabrany, H., Agustina, R., Dartanto, T., Sitompul, R., Susiloretni, K. A., Suparmi, Achadi, E. L., Taher, A., ... Khusun, H. (2019). Universal health coverage in Indonesia: Concept, progress, and challenges. The Lancet, 393(10166), 75–102. https://doi.org/10.1016/S0140-6736(18)31647-7

Akwaowo, C. D., Umoh, I., Motilewa, O., Akpan, B., Umoh, E., Frank, E., Nna, E., Okeke, U., & Onwujekwe, O. E. (2021). Willingness to Pay for a Contributory Social Health Insurance Scheme: A Survey of Rural Residents in Akwa Ibom State, Nigeria. Frontiers in Public Health, 9, 654362. https://doi.org/10.3389/fpubh.2021.654362

Asenso-Okyere, W. K., Osei-Akoto, I., Anum, A., & Appiah, E. N. (1997). Willingness to pay for health insurance in a developing economy. A pilot study of the informal sector of Ghana using contingent valuation. Health Policy, 42(3), 223–237. https://doi.org/10.1016/S0168-8510(97)00069-9

Ayre, C., & Scally, A. J. (2014). Critical Values for Lawshe's Content Validity Ratio: Revisiting the Original Methods of Calculation. Measurement and Evaluation in Counseling and Development, 47(1), 79–86. https://doi.org/10.1177/0748175613513808

Baker, C. L., Bruno, M., Grant, L., Johnson, C., Bennett, B., Brohan, E., & Emir, B. (2017). Content Validity of a Willingness to Quit Tool for Use with Current Smokers in Clinical Practice. Advances in Therapy, 34(10), 2295–2306. https://doi.org/10.1007/s12325-017-0611-y

Bourke, S., McColl, E., Shaw, P., & Gibson, G. (2004). Validation of quality of life instruments in ALS. Amyotrophic Lateral Sclerosis and

Other Motor Neuron Disorders, 5(1), 55–60. https://doi.org/10.1080/14660820310016066

Breidert, C. (2005). *Estimation of willingness-to-pay. Theory, measurement, and application.* (Doctoral thesis, WU Vienna).

Chanie, M. G., & Ewunetie, G. E. (2020). Determinants of enrollment in community based health insurance among Households in Tach-Armachiho Woreda, North Gondar, Ethiopia, 2019. PLOS ONE, 15(8), e0236027. https://doi.org/10.1371/journal.pone.0236027

Chiwaridzo, M., Chikasha, T. N., Naidoo, N., Dambi, J. M., Tadyanemhandu, C., Munambah, N., & Chizanga, P. T. (2017). Content validity and test-retest reliability of a low back pain questionnaire in Zimbabwean adolescents. Archives of Physiotherapy, 7(1), 3. https://doi.org/10.1186/s40945-017-0031-y

Cook, D. A., & Beckman, T. J. (2006). Current Concepts in Validity and Reliability for Psychometric Instruments: Theory and Application. The American Journal of Medicine, 119(2), 166.e7-166.e16. https://doi.org/10.1016/j.amjmed.2005.10.036

Darmawan, K. H., Satibi, S., & Kristina, S. A. (2019). Willingness To Pay for Social Health Insurance and Related Factors Among Population in Yogyakarta Province, Indonesia. International Medical JournaL, 22(April), S787. https://doi.org/10.1016/j.jval.2019.09.2053

Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. Applied Nursing Research, 5(4), 194–197. https://doi.org/10.1016/S0897-1897(05)80008-4

Deksisa, A., Abdo, M., Mohamed, E., Tolesa, D., Garoma, S., Zewdie, A., Lami, M., Irena, D., Abdena, D., & Lemi, H. (2020). Willingness of community based health insurance uptake and associated factors among urban residents of Oromia regional state, Oromia, Ethiopia, a cross-sectional study. BMC Health Services Research, 20(1), 864. https://doi.org/10.1186/s12913-020-05583-x

Fallahzadeh, H., Etesam, F., & Asgarian, F. S. (2015). Validity and reliability related to the Persian version of the Children's Sleep Habits

Questionnaire: Validity of the Persian CHSQ. Sleep and Biological Rhythms, 13(3), 271–278. https://doi.org/10.1111/sbr.12114

Farzanegan, R., Farzanegan, B., Alehashem, M., Zangi, M., Niakan Kalhori, S. R., Sheikhy, K., Emami, H., & Behgam Shadmehr, M. (2017). Item Selection and Content Validity of the Risk Factors of Post-Intubation Tracheal Stenosis Observation Questionnaire for ICU-Admitted Patients. Tanaffos, 16(1), 22–33.

Feng, J., Song, H., & Wang, Z. (2020). The elderly's response to a patient cost-sharing policy in health insurance: Evidence from China. Journal of Economic Behavior & Organization, 169, 189–207. https://doi.org/10.1016/j.jebo.2019.11.009

Foreit, J. R., & Foreit, K. G. F. (2003). The reliability and validity of willingness to pay surveys for reproductive health pricing decisions in developing countries. Health Policy, 63(1), 37–47. https://doi.org/10.1016/S0168-8510(02)00039-8

Gidey, M. T., Gebretekle, G. B., Hogan, M.-E., & Fenta, T. G. (2019). Willingness to pay for social health insurance and its determinants among public servants in Mekelle City, Northern Ethiopia: A mixed methods study. Cost Effectiveness and Resource Allocation, 17(1), 2. https://doi.org/10.1186/s12962-019-0171-x

Grant, J. S., & Davis, L. L. (1997). Selection and use of content experts for instrument development. Research in Nursing & Health, 20(3), 269–274. https://doi.org/10.1002/(SICI)1098-240X(199706)20:3<269::AID-NUR9>3.0.CO;2-G

Haile, M., Ololo, S., & Megersa, B. (2014). Willingness to join community-based health insurance among rural households of Debub Bench District, Bench Maji Zone, Southwest Ethiopia. BMC Public Health, 14(1), 591. https://doi.org/10.1186/1471-2458-14-591

Hossein, Z., & Gerard, A. (2013). Trends in cost sharing among selected high income countries—2000–2010. Health Policy, 112(1–2), 35–44. https://doi.org/10.1016/j.healthpol.2013.05.020

Juengst, S. B., Terhorst, L., Dicianno, B. E., Niemeier, J. P., & Wagner, A. K. (2019). Development and content validity of the behavioral

assessment screening tool (BAST β). Disability and Rehabilitation, 41(10), 1200–1206. https://doi.org/10.1080/09638288.2017.1423403

Kado, A., Merga, B. T., Adem, H. A., Dessie, Y., & Geda, B. (2020). Willingness to Pay for Community-Based Health Insurance Scheme and Associated Factors Among Rural Communities in Gemmachis District, Eastern Ethiopia. ClinicoEconomics and Outcomes Research: CEOR, 12, 609–618. https://doi.org/10.2147/CEOR.S266497

Kawakami, N., Inoue, A., Tsuchiya, M., Watanabe, K., Imamura, K., Iida, M., & Nishi, D. (2020). Construct validity and test-retest reliability of the World Mental Health Japan version of the World Health Organization Health and Work Performance Questionnaire Short Version: A preliminary study. Industrial Health, 58(4), 375–387. https://doi.org/10.2486/indhealth.2019-0090

Koohi Rostamkalaee, Z., Jafari, M., & Gorji, H. A. (2022). A systematic review of strategies used for controlling consumer moral hazard in health systems. BMC Health Services Research, 22(1), 1260. https://doi.org/10.1186/s12913-022-08613-y

Lau, A. S. Y., Yusoff, M. S. B., Lee, Y.-Y., Choi, S.-B., Xiao, J.-Z., & Liong, M.-T. (2018). Development and validation of a Chinese translated questionnaire: A single simultaneous tool for assessing gastrointestinal and upper respiratory tract related illnesses in pre-school children. Journal of Taibah University Medical Sciences, 13(2), 135–141. https://doi.org/10.1016/j.jtumed.2017.11.003

Lavoie Smith, E., Haupt, R., Kelly, J., Lee, D., Kanzawa-Lee, G., Knoerl, R., Bridges, C., Alberti, P., Prasertsri, N., & Donohoe, C. (2017). The Content Validity of a Chemotherapy-Induced Peripheral Neuropathy Patient-Reported Outcome Measure. Oncology Nursing Forum, 44(5), 580–588. https://doi.org/10.1188/17.ONF.580-588

Lawshe, C. H. (1975). A Quantitative Approach to Content Validity. Personnel Psychology, 28(4), 563–575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

Lee, J. L., Chung, Y., Waters, E., & Vedam, H. (2020). The Epworth sleepiness scale: Reliably unreliable in a sleep clinic population. Journal of Sleep Research, 29(5). https://doi.org/10.1111/jsr.13019

Lynn, M. R. (1986). Determination and Quantification Of Content Validity: Nursing Research, 35(6), 382–386. https://doi.org/10.1097/00006199-198611000-00017

Mekonne, A., Seifu, B., Hailu, C., & Atomsa, A. (2020). Willingness to Pay for Social Health Insurance and Associated Factors among Health Care Providers in Addis Ababa, Ethiopia. BioMed Research International, 2020, 1–7. https://doi.org/10.1155/2020/8412957

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. American Psychologist, 50(9), 741–749. https://doi.org/10.1037/0003-066X.50.9.741

Minyihun, A., Gebregziabher, M. G., & Gelaw, Y. A. (2019). Willingness to pay for community-based health insurance and associated factors among rural households of Bugna District, Northeast Ethiopia. BMC Research Notes, 12(1), 1–7. https://doi.org/10.1186/s13104-019-4091-9

Moghaddam, N. M., Jame, S. Z. B., Rafiei, S., Sarem, A. A., Ghamchili, A., & Shafii, M. (2019). Managerial competencies of head nurses: A model and assessment tool. British Journal of Nursing, 28(1), 30–37. https://doi.org/10.12968/bjon.2019.28.1.30

Mohamad Marzuki, M. F., Yaacob, N. A., & Yaacob, N. M. (2018). Translation, Cross-Cultural Adaptation, and Validation of the Malay Version of the System Usability Scale Questionnaire for the Assessment of Mobile Apps. JMIR Human Factors, 5(2), e10308. https://doi.org/10.2196/10308

Moshki, M., Khajavi, A., Sadeghi-Bazargani, H., Vahedi, S., & Pour-Doulati, S. (2020). Developing Pedestrians' Red-light Violation Behavior Questionnaire (PRVBQ); Assessment of Content Validity and Reliability. Bulletin of Emergency and Trauma, 8(2), 98–106.

https://doi.org/10.30476/BEAT.2020.4644 9

Nguyen, H. T., & Connelly, L. B. (2017). *Cost-sharing in health insurance and its impact in a developing country– Evidence from a quasi-natural experiment*. Retrieved from https://mpra.ub.uni-muenchen.de/76399/ (accessed on 30 August 2019).

Nosratnejad, S., Rashidian, A., Mehrara, M., Akbari Sari, A., Mahdavi, G., & Moeini, M. (2014). Willingness to pay for the social health insurance in Iran. Global Journal of Health Science, 6(5), 154–163. https://doi.org/10.5539/gjhs.v6n5p154

Nugraheni, D. A., Satibi, S., Kristina, S. A., & Puspandari, D. A. (2022). Factors Associated with Willingness to Pay for Cost-Sharing under Universal Health Coverage Scheme in Yogyakarta, Indonesia: A Cross-Sectional Survey. International Journal of Environmental Research and Public Health, 19(22), 15017. https://doi.org/10.3390/ijerph192215017

Padilla, J.-L., & Benítez, I. (2014). Validity evidence based on response processes. Psicothema, 26.1, 136–144. https://doi.org/10.7334/psicothema2013. 259

Paiva, C. E., Carneseca, E. C., Barroso, E. M., De Camargos, M. G., Alfano, A. C. C., Rugno, F. C., & Paiva, B. S. R. (2014). Further evaluation of the EORTC QLQ-C30 psychometric properties in a large Brazilian cancer patient cohort as a function of their educational status. Supportive Care in Cancer. https://doi.org/10.1007/s00520-014-2206-3

Park, E., & Choi, S. (2020). Who Benefits from the Fixed Copayment of Medical and Pharmaceutical Expenditure among the Korean Elderly? International Journal of Environmental Research and Public Health, 17(21), 8118. https://doi.org/10.3390/ijerph17218118

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? critique and recommendations. Research in Nursing & Health, 29(5), 489–497. https://doi.org/10.1002/nur.20147

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Research in Nursing & Health, 30(4), 459–467. https://doi.org/10.1002/nur.20199

Rascati, K. (2014). *Essentials of Pharmacoeconomics*. 2nd ed. Philadelphia: Lippincott Williams & Wilkins. pp. 109.

Remler, D. K., & Greene, J. (2009). Cost-Sharing: A Blunt Instrument. Annual Review of Public Health, 30(1), 293–311. https://doi.org/10.1146/annurev.publhealt h.29.020907.090804

Rozgonyi, R., Dombi, I., Janszky, J., Kovács, N., & Faludi, B. (2021). Low test–retest reliability of the Epworth Sleepiness Scale within a substantial short time frame. Journal of Sleep Research, 30(4), e13277. https://doi.org/10.1111/jsr.13277

Saber, A., Tabatabaei, S. M., Akasheh, G., Sehat, M., Zanjani, Z., & Larijani, B. (2017). Face and Content Validity of the MacArthur Competence Assessment Tool for the Treatment of Iranian Patients. International Journal of Preventive Medicine, 8, 75. https://doi.org/10.4103/ijpvm.IJPVM_367_ 16

Salampessy, B. H., Alblas, M. M., Portrait, F. R. M., Koolman, X., & Van Der Hijden, E. J. E. (2018). The effect of cost-sharing design characteristics on use of health care recommended by the treating physician; A discrete choice experiment 11 Medical and Health Sciences 1117 Public Health and Health Services 14 Economics 1402 Applied Economics. BMC Health Services Research, 18(1), 1–14. https://doi.org/10.1186/s12913-018-3598-4

Sangoseni, O., Hellman, M., & Hill, C. (2013). Development and Validation of a Questionnaire to Assess the Effect of Online Learning on Behaviors, Attitudes, and Clinical Practices of Physical Therapists in the United States Regarding Evidenced-based Clinical Practice. Internet Journal of Allied Health Sciences and Practice. https://doi.org/10.46743/1540-580X/2013.1439

Seaman, K. L., Sanfilippo, F. M., Roughead, E. E., Bulsara, M. K., Kemp-Casey, A., Bulsara, C., Watts, G. F., & Preen, D. (2017). Impact of consumer copayments for subsidised medicines on health services use and outcomes: A protocol using linked administrative data from Western Australia. BMJ Open, 7(6), e013691. https://doi.org/10.1136/bmjopen-2016-013691

Thi Thuy Nga, N., FitzGerald, G., & Dunne, M. P. (2018). Family-Based Social Health Insurance for Informal Workers in Vietnam: Willingness to Pay and Its Determinants. Asia-Pacific Journal of Public Health, 30(6), 512–520. https://doi.org/10.1177/10105395187997 85

Vartanian, J. G., Carvalho, A. L., Yueh, B., Furia, C. L. B., Toyota, J., McDowell, J. A., Weymuller, E. A., & Kowalski, L. P. (2006). Brazilian–Portuguese validation of the University of Washington Quality of Life Questionnaire for patients with head and neck cancer. Head & Neck, 28(12), 1115–1121. https://doi.org/10.1002/hed.20464

Vaz, S., Falkmer, T., Passmore, A. E., Parsons, R., & Andreou, P. (2013). The Case for Using the Repeatability Coefficient When Calculating Test–Retest Reliability. PLoS ONE, 8(9), e73990. https://doi.org/10.1371/journal.pone.0073 990

World Health Organization. (2010). *The World Health Report: Health Systems Financing: The Path to Universal Coverage*. World Health Organization. Retrieved from https://www.who.int/publications/i/item/ 9789241564021 (accessed on 1 August 2022).

Yusoff, M. S. B. (2019). ABC of Content Validation and Content Validity Index Calculation. Education in Medicine Journal, 11(2), 49–54. https://doi.org/10.21315/eimj2019.11.2.6

Zamanzadeh, V., Rassouli, M., Abbaszadeh, A., Majd, H. A., Nikanfar, A., & Ghahramanian, A. (2014). Details of content validity and objectifying it in instrument development. Nursing Practice Today, 1(3), 163–171.