BMJ Open

Development and validation of the health behavioural intention among at-risk smokers to prevent nasopharyngeal cancer in Sarawak, Malaysia based on the Health Belief Model

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-057552
Article Type:	Original research
Date Submitted by the Author:	29-Sep-2021
Complete List of Authors:	Kueh, Martin; RCSI & UCD Malaysia Campus, Department of Public Health Medicine Rahim, Fairuz Fadzilah; RCSI & UCD Malaysia Campus, Department of Public Health Medicine Rashid, Abdul; RCSI & UCD Malaysia Campus, Public Health Medicine
Keywords:	PUBLIC HEALTH, Head & neck tumours < ONCOLOGY, Epidemiology < ONCOLOGY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Development and validation of the health behavioural intention among at-risk smokers to prevent nasopharyngeal cancer in Sarawak, Malaysia based on the Health Belief Model

Martin Kueh[1], Fairuz Fadzilah Rahim[1] & Abdul Rashid[1]

[1] Department of Public Health Medicine, RCSI & UCD Malaysia Campus, No.4, Jalan Sepoy Lines, 10450 George Town, Penang Malaysia

Correspondence: Martin Kueh, Department of Public Health Medicine, RCSI & UCD Malaysia Campus, No.4, Jalan Sepoy Lines, 10450 George Town, Penang Malaysia. Tel: 6016-879 2269. Email: martin1kueh@hotmail.com

Abstract

Objective Since lifestyle-induced NPC should not be disregarded, this study serves to develop and validate a questionnaire that aims to predict the health behaviour intentions of at-risk smokers in Sarawak, Malaysia using Health Belief Model (HBM) as the conceptual framework.

Design Prospective validation cross sectional study

Setting Urban and suburban areas in Sarawak, Malaysia.

Participants The preliminary items of the instrument were developed through a literature review. The instrument was translated into Malay version using forward-backward method before conducting content validity through a panel of 10 experts. Face validity was examined both quantitatively and qualitatively by 10 local smokers. The construct validity of the instrument was evaluated through exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). A total of 100 local smokers participated in Phase 1 for EFA, while 171 local smokers participated in Phase 2 for CFA. Internal consistency was measured using Cronbach's alpha coefficients to evaluate the reliability.

Results In the exploratory stage, the factor loading of each item remained within the acceptable threshold. The final revised CFA yielded appropriate fit of the seven-factor model with the following model fit indexes: Chi Square: 641.705; df= 500; P< 0.001; CFI = 0.953; TLI: 0.948; RMSEA= 0.041. Satisfactory convergent validity and divergent validity were shown, with the exception of one pairwise construct. Internal reliability of these scales was above the desirable threshold, with Cronbach's alpha coefficients ranging from 0.705 to 0.864 and 0.838 to 0.889 for Phase 1 and 2, respectively.

Conclusions The study substantiated the instrument to be valid and reliable for predicting smokers' health behavioural intention to reduce cancer risk. The instrument is made up of 34 items, categorised into two sections, six HBM constructs and health behavioural intention. The instrument can be utilised for other smoking-related cancers in different at-risk population.

Keywords: tobacco smoking, cancer health promotion, nasopharyngeal cancer, health belief model, development and validation, Malaysia

Article Summary

Strengths and limitations of this study

1. This study established a novel instrument for assessing at-risk smokers' health behavioural intention to reduce NPC risk based on a well-known framework in a series of systematic validation stages, which potentially can be applied for other smoking-related cancers.

2. Face validity was undertaken both qualitatively and quantitatively to sufficiently reflect the demographic during the assessments of psychological constructs, and the validation was markedly aided by experts' evaluations.

3. This study was conducted in two phases, involving both urban and suburban local smokers, to examine concept validity, convergent validity, and divergent validity.

4. The study's generalizability may be limited by the smokers' cultural perspective, hence, further studies on smokers from different cultures will be needed to assess the instrument's psychometric properties.

Introduction

According to the World Health Organization, tobacco smoking is a public health concern that accounts for over 8 million deaths per year and is the leading avoidable cause of illness, disability, and death globally [1]. Annually, the exposure to smoking is associated with 2.4 million deaths from cancer throughout the world [2]. The Surgeon General's report on the relationship between smoking and cancer was a watershed moment in public health towards tobacco's adverse effects on human health, followed by subsequent discovery that tobacco smoke comprises approximately 7,000 com-pounds, 72 of which have been identified as carcinogenic [3, 4]. Tobacco use is now causally associated to at least 20 cancer types, which also appears to have a wide-ranging immediate and long-term health benefits accompany cessation [2]. However, the harmful consequences of tobacco smoking are widely neglected or underestimated, despite the fact that it remains a significant public health hazard with the impoverished, and marginalised, as well as those in developing nations, bearing a disproportionate share of the burden [5].

Cancer is a leading cause of death as well as a major obstacle to improving lifespan in every country [6]. There is an approximately 1% reduction in overall cancer mortality rate across both sexes in both high- and low-income countries [7]. However, this is mainly a result of positive trends in the most common cancers. It fails to account for variations in the frequency and distribution of etiologic aspects such as socioeconomic, geographical, genetic, biological, ethnic, social, physical factors, as well as disparities across cancer types [7, 8]. For instance, nasopharyngeal cancer (NPC) is uncommon but unique among head and neck cancers and is a cancer disparity with its own distinct epidemiological & risk factor profiles. To reinforce the claim, the Global data from the World Health Organization illustrates poorer outcomes of NPC in endemic areas like Southeast Asia with an unbalanced global burden of 67% [9].

In Malaysia, NPC is a nationwide public health concern and the 5th most leading form of cancer, amounted to 4,597 new diagnoses of NPC for the 2012-2016 periods. The re-cent report from the Malaysian National Cancer Registry reports that the lifetime risk of developing NPC among men and women are 1 in 175 and 1 in 482, respectively [10]. Strikingly, there is a substantial geographical variance across the country, with Sarawak exhibiting a higher prevalence rate of NPC. A previous study

has shown a significant high age-standardised rates in males (13.5/100,000, 95% Confidence Interval = 12.2 – 15.0) and females (6.2/100,000, 95% CI= 5.7-6.7) by which the local at-risk ethnic groups including Bidayuh, Chinese, Iban, Malays and Melanau were collectively ranked top globally. In particular, the risk among the Bidayuh ethnic population, which is a native indigenous group, exceeds the general population of males and females in Sarawak by 2.3 times and 1.9 times respectively [11].

This trend was ascribed to potential risk factors, which include Epstein-Barr virus, genetic susceptibility, consumption of food with nitrous compounds and volatile nitrosamines, and complex interaction with environmental factors [12]. Among many risk factors that are associated with NPC, tobacco smoking is the most important modifiable cause of severity of NPC [13, 14]. A past study analyses 32 epidemiologic studies (28 case control studies and 4 cohort studies) regarding tobacco smoking and NPC from 1979 to 2011 and reports extensively that tobacco-correlated NPC cases are 60% higher compared to non-smokers [15]. Further, National Health Morbidity Survey 2015 shows that the prevalence of tobacco smoking among the population in Sarawak is 25.4%. The native indigenous male (61.2%) and female smokers (10.7%) in Sarawak are among the highest nationwide [16]. Rahman et al indicate that an average number of tobaccos smoked is 13.6 cigarettes per day in Sarawak [17]. This lifestyle-induced NPC should not be disregarded, and it necessitates a preventative strategy centred on modifying health risk behaviours.

Despite the robust establishment of cumulative impacts of tobacco smoking on the risk of cancer, there is still paucity of local studies and research effort to discover critical areas and serve as a benchmark evaluation for effectiveness of comprehensive strategies to promote cancer prevention among smokers. Therefore, given the significance of this topic, it is critical to create a questionnaire that focuses on behavioural factors and is customised to the interests of local at-risk smokers. This study serves to develop and validate a health behaviour model (HBM)-based questionnaire that aims to predict the health behaviour intentions of these smokers and their perspective and motivation towards quit smoking. A health intervention targeting NPC in this population may be more effective due to the awareness of this population regarding their increased susceptibility to NPC and may also benefit from being informed by this questionnaire.

Conceptual Framework

The Health Belief Model (HBM) is the underpinning conceptual framework for this study to focus on psychological variables for predicting health behavioural intention. Developed in the 1950s by social psychologists at the United States Public Health Ser-vice, the Health Belief Model (HBM) is currently one of the most extensively used cognitive model and theoretical framework to help researchers understand and predict health behaviours in the population and ultimately to inform & direct health promotion and interventions [18, 19]. A large volume of past studies utilized HBM to examine health-promoting behaviours of prevention for different cancers in the context of both developed and developing countries, for example, the United States, Iran and Ethiopia [20-22]. The HBM is a value-expectancy theory, based on the hypothesis of Lewin et al that highlights the influence of two variables on behaviour: 1) the value that a person places on the outcome of the behaviour and 2) the person's perception of how likely the behaviour will lead to that outcome, in the event of illness (19). Having evolved over the past decades, HBM currently consists of 6 elements: i) perceived susceptibility, ii) perceived severity, iii) perceived benefits, iv) perceived barriers, v) cues to action, and vi) self-efficacy [23].

The first 4 refer to a person's subjective perceptions regarding 1) his/her risk of getting the disease; 2) how severe the consequences are of getting the disease; 3) the benefits from performing a health behaviour in preventing, curing, or managing the disease; and 4) obstacles to that health behaviour, e.g. financial and time costs, side effects, and so on (23). "Cue to action" is the stimulus, which may be internal (e.g. physical sensations) or external (e.g. friends with the disease, social media), that is required for that health behaviour to occur, and "self-efficacy" refers to the person's confidence how capable he/she is to successfully undertake that health behaviour [19, 23].

The following conceptual framework depicts the proposed causal relationship between HBM constructs and health behavioural intention (Figure 1):

Methods

Study design and setting

A cross-sectional study was conducted in Sarawak, Malaysia in two phases: Phase 1 from October 2020 to January 2021 and Phase 2 from January to April in 2021. Sarawak is the largest Malaysian state on the island of Borneo with a population of more than 2.6 million, made up of 26 different ethnic groups. Sarawak is divided into twelve divisions, each of which is further divided into districts and sub districts.

The sample population of Phase 1 was residents residing in urban and suburban areas in Miri, Sarawak. Phase 2 was mainly involved by residents residing in Bintulu, Kuala Baram (a federal constituency in Miri Division), and remote rural areas in Miri. They were mainly local employees working in the agricultural industry. Data were collected by eight trained research assistants. All research assistants were given a crash course in research aims, methodology and data collection, as well as a trial run to simulate real-world situations. Prior to distribution of questionnaires, participants were briefed regarding the objective and methodology of the study, as well as the benefits and risks. Involvement in this sample was entirely voluntary and did not pose any potential threat. Upon clarifying the study details, informed consents were obtained by eligible participants. The anonymity of the respondents' details was assured. The research assistants provided clarifications to smokers who requested assistance, and their replies were meticulously documented. The methodology of the study and data collection are recorded precisely and accurately throughout the process of the research.

The inclusion criteria were: 1) 18 years old and above, 2) smoked for at least a year, and 3) is a Sarawakian. Participants who did not consent to join, were pregnant, or smoked e-cigarettes only were excluded from this study.

Patient and Public Involvement

No patient involved

Instrument development and validation

The questionnaire is self-administered based on prior validated studies. Search was conducted using the National Library of Medicine (PubMed), Google Scholar and Cochrane Library databases by exploring various keywords: health belief model, nasopharyngeal cancer, cancer, smokers, smoking behaviours/habit, and questionnaire/tools/instruments. Questionnaires and prior literatures are adapted to examine primarily, and explicitly on smokers about NPC prevention using HBM. The initial questionnaire consisted of 41 items.

The questionnaire was developed in English and translated to Malay by two local fluent bilingual translators. An experienced researcher, whose mother tongue is Malay, compared the Malay version questionnaire to the English questionnaire. A 'back-translation' approach to English was taken independently by another two bilingual translators based on the Beaton-recommended guidelines [24].

To determine the content validity, 10 healthcare professionals including public health experts, hospital directors, and health officers were invited to evaluate the survey instrument. According to Lawshe's model, a questionnaire was designed and organized to assist and allow panellists to express clearly

their views on the importance of include different components in a model. Experts received attachments of the questionnaire via email, which was graded on a three-point scale: essential, useful but not essential, and not essential. Based on Lawshe's table, items with a CVR value greater than 0.62 were kept for this study [25]. To minimise any ambiguity, the experts evaluated each item's accuracy, phrasing and grammar as well as their relevance to the construct. Modifications were made for subsequent analysis based on experts' comments and suggestions. Overall, 7 items were deleted, 2 items were rephrased and 1 item was allocated into different construct.

Face validity was evaluated through a pilot study by 10 local smokers of different ethnic background, both qualitatively and quantitatively. Smokers who participated in the pilot study were exempted from the main study. The pilot study is conducted on a small-scale basis to ascertain the feasibility of the proposed larger study [26]. In the qualitative stage, minor revisions were made to better suit a linguistically and culturally diverse context. Based on local smokers' perspective, an item oriented towards a person's faith or spirituality to own health should also be included. Upon consensus from the researchers, we decided to add in one item, 'I think getting nose and throat cancer is my destiny and quitting smoking will not change it'. In the quantitative stage, a 5-point Likert scale Impact Score, ranging from 'extremely important' to 'least important', was measured. Once the smokers completed the questionnaire, the first step was to determine the 'Frequency' by counting the proportion of individuals who rated 4 or 5 on item im-portance. The second step was to determine the 'Importance' by calculating the mean importance score of each item. Finally, each item's Impact Score is computed using the following formula: Impact Score = Frequency (Proportion) x Importance [27]. If an item's Impact Score is equal to or more than 1.5, it is deemed suitable and kept for further evaluation [28]. On the whole, the impact score for each item ranged from 1.7 to 4.6.

Subsequently, in the main study, to determine the construct validity, exploratory and confirmatory factor analysis were performed in two periods. In the first period, EFA (n = 100) was used to determine the number of latent factors or the relationships between the common factors. The model was later adjusted in the second period with CFA (n = 171) via structural equations using AMOS. CFA confirmed the overall fit of the model and indicated that the measures were in acceptable range [29]. Convergent validity and discriminant validity were also carried out in Phase 2.

Ethical considerations:

This study was approved by RCSI & UCD Malaysia Campus (RUMC) Institutional Research and Ethics Committee (Approval no. JPEC 20 0027). The informed consent of all participants was obtained voluntarily, and the data was kept entirely confidentially.

Data management:

The data was processed in Microsoft Excel before being analysed in SPSS 26.0 and AMOS 23.0. Socio-demographic characteristics are presented as number and percent-age distribution. Cronbach's alpha coefficients was used to determine internal consistency. Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) were used to test the construct validity of each construct. Specifically, EFA was evaluated in Phase 1 and CFA in Phase 2. A p-value below 0.05 was deemed to be statistically significant.

EFA was performed in Phase 1 to reveal the fundamental structure of a large set of variables [29]. The factors were extracted using principal component analysis with a varimax rotation. As for sampling adequacy and item-checking, the Kai-ser-Mayer-Olkin (KMO) > 0.6 and Bartlett's test for sphericity ($p < 0.05$) were used. All loading factors below 0.3 were excluded from the constructs [30].

CFA was performed in Phase 2 to assess the data integrity and the structural model [29]. The acceptable level of standardised factor loading was set at 0.5 and above to ensure a satisfactory association between items and corresponding factors [31]. Different fit indices were utilised to estimate the model fit, for instance a comparative fit index (CFI) of > 0.90, Tucker-Lewis Index (TLI) > 0.90 and root mean square error of approximation (RMSEA) < 0.05 [32]. Additionally, convergent validity and discriminant validity were evaluated based on the composite reliability (CR) and average variance extracted (AVE).

Results:

A total of 100 and 171 smokers participated in Phase 1 and Phase 2 of the study, respectively. The majority of the participants were males (Phase 1 86.0%; Phase 2 89.5%) and were in the 30-39 age group (Phase 1 40.0%; Phase 2 40.9%). A little over a third of the participants smoked for more than 10 years for both Phase 1 (37%) and Phase 2 (40.4%). In Phase 1, most of the participants were Iban (34%), followed by Malay (19%), Chinese (18%), Others (13%), Melanau (9%) and Bidayuh (7%). Most of the participants in Phase 2 were Chinese (33.9%), followed by Iban (24.6%), Malay (12.3%), Bidayuh (9.9%), Melanau (9.9%) and Others (9.4%). Details of the smokers' demographic information are presented in Table 1.

Table 1: Demographic characteristics of participants

	Phase 1 (n=100)		Phase 2 (n=171)	
	Number	Percentage (%)	Number	Percentage (%)
Age				
18-29	34	34.0	46	26.9
30-39	40	40.0	70	40.9
40-49	14	14.0	43	25.1
50-64	12	12.0	10	5.8
65 and above	0	0.0	2	1.2
Gender				
Male	86	86.0	153	89.5
Female	14	14.0	18	10.5
Ethnic groups				
Malay	19	19.0	21	12.3
Chinese	18	18.0	58	33.9
Bidayuh	7	7.0	17	9.9
Iban	34	34.0	42	24.6
Melanau	9	9.0	17	9.9
Others	13	13.0	16	9.4
Years of smoking				
1-5 years	32	32.0	44	25.7
6-10 years	31	31.0	58	33.9
More than 10 years	37	37.0	69	40.4

In Phase 1, Kaiser-Meyer-Olkin Measure of Sampling Adequacy was 0.697 and Barlett's test of the sphericity was significant (x2 = 1,746, p-value < 0.001). EFA was conducted to analyse the factor structure with principal component analysis with a varimax rotation. A decision was made to go for 7-factor structures since there is clarity of 7 constructs. The EFA found that 7 variables had eigenvalues larger than Kaiser's threshold of 1 and explained 63.0% of the variance when combined. Factor loadings of higher than 0.3 were found in all the items. Four items had cross loading with values greater than 0.3, which are PBar5, HBI2, HBI3 & HBI4. All items remain because the contents of the items were regarded as relevant based on the decision and judgment of the researchers. Table 2 shows the EFA with total items and the factor loading of each construct for the 7-factor model.

Table 2: Result of Exploratory Factor Analysis in Phase 1 (n = 100)

Constructs	Items	Component						
		1	2	3	4	5	6	7
Perceived Susceptibility	PSus1			0.665				
	PSus2			0.750				
	PSus3			0.810				
	PSus4			0.548				
	PSus5			0.736				
Perceived Severity	PSev1						0.558	
	PSev2						0.720	
	PSev3						0.744	
	PSev4						0.843	
	PSev5						0.730	
Perceived Benefit	PBen1	0.767						
	PBen2	0.765						
	PBen3	0.763						
	PBen4	0.792						
Perceived Barrier	PBar1					0.594		
	PBar2					0.725		
	PBar3					0.658		
	PBar4					0.709		
	PBar5					0.590		
	PBar6					0.617		
Cue to action	CUE1				0.721			
	CUE2				0.584			
	CUE3				0.736			
	CUE4				0.723			
	CUE5				0.745			
Self-efficacy	EFF1		0.721					
	EFF2		0.730					
	EFF3		0.770					
	EFF4		0.868					
	EFF5		0.625					
Health Behavioural	HBI1							0.681
	HBI2							0.307

Intention	HBI3							0.733
	HBI4							0.596
Rotation Sums of Squared Loading								
Total		3.469	3.371	3.246	3.072	3.071	3.012	2.185
Percentage of Variance		10.203	9.913	9.548	9.034	9.032	8.859	6.425
Cumulative percentage		10.203	20.116	29.664	38.698	47.730	56.588	63.014

CFA was performed in Phase 2 to assess whether the seven-factor model indicated by the EFA could sufficiently represent the data. The items in their respective constructs were loaded between 0.586 and 0.898 (Table 3). For the model's fitness to increase, items with less than 0.6 and a MI more than 10 should be eliminated. Despite this, they were kept owing to the essential for the conceptual framework. Before arriving at the final model, 6 pairs of correlated errors were added to improve robustness. The resulting model is suitable for testing, as evidenced by the following model fit indexes: Chi Square: 641.705; df= 500; P< 0.001; CFI = 0.953; TLI: 0.948; RMSEA= 0.041 (90% CI 0.031 to 0.050).

Table 3: Result of Confirmatory Factor Analysis in Phase 2 (n = 171)

Constructs	Items	Factor Loadings	AVE	CR
Perceived Susceptibility	PSus1	0.898	0.577	0.871
	PSus2	0.819		
	PSus3	0.739		
	PSus4	0.685		
	PSus5	0.683		
Perceived Severity	PSev1	0.680	0.597	0.881
	PSev2	0.675		
	PSev3	0.806		
	PSev4	0.867		
	PSev5	0.766		
Perceived Benefit	PBen1	0.752	0.603	0.858
	PBen2	0.812		
	PBen3	0.740		
	PBen4	0.733		
Perceived Barrier	PBar1	0.586	0.572	0.888

	PBar2	0.670		
	PBar3	0.626		
	PBar4	0.886		
	PBar5	0.878		
	PBar6	0.770		
	CUE1	0.646		
	CUE2	0.721		
	CUE3	0.727		
	CUE4	0.767		
Cue to action	CUE5	0.723	0.512	0.839
	EFF1	0.645		
	EFF2	0.694		
	EFF3	0.827		
	EFF4	0.859		
Self-efficacy	EFF5	0.744	0.574	0.869
	HBI1	0.654		
	HBI2	0.836		
Health Behavioural Intention	HBI3	0.867		
	HBI4	0.773	0.617	0.864

The AVE and CR values, which are listed in Table 3, were obtained after the structural model's fit was investigated to check if the items were loaded up appropriately. The AVE readings were all over the cut-off value of 0.5, ranging from 0.512 to 0.617. The CR values were all over the cut-off value of 0.7, ranging from 0.839 to 0.888. All seven constructs featured sufficient convergent validity.

Discriminant validity was evaluated using Fornell-Larcker criteria by comparing the squared correlations and AVE scores for each of the pairwise constructs (33). With the exception of Perceived Benefit < — > Cue to action, all paired constructs have shown established discriminant validity (Table 4).

Table 4: Result of Discriminant Validity in Phase 2 (n = 171)

	Factor Correlation	Correlation Squared	Discriminant validity
Perceived Susceptibility<-->Perceived Severity	.312	0.097	Established
Perceived Susceptibility <-->Perceived Benefit	.260	0.068	Established
Perceived Susceptibility <-->Perceived Barrier	.204	0.042	Established
Perceived Susceptibility <-->Cue to Action	.172	0.030	Established
Perceived Susceptibility <-->Self-Efficacy	.236	0.056	Established
Perceived Severity <--> Perceived Benefit	.575	0.331	Established
Perceived Severity <--> Perceived Barrier	-.042	0.176	Established
Perceived Severity <--> Cue to Action	.238	0.057	Established
Perceived Severity <--> Self-Efficacy	.257	0.066	Established
Perceived Benefit <--> Perceived Barrier	-.085	0.007	Established

Perceived Benefit <--> Cue to Action	.826	0.682	Not Established
Perceived Benefit <--> Self-Efficacy	.349	0.122	Established
Perceived Barrier <--> Cue to Action	-.001	0.000	Established
Perceived Barrier <--> Self-Efficacy	-.157	0.025	Established
Cue to Action <--> Self-Efficacy	.600	0.360	Established

Internal consistency was deemed to be acceptable if the Cronbach's alpha coefficients were more than 0.7. According to reliability analyses, Cronbach α of the Perceived Susceptibility, Perceived Severity, Perceived Benefits, Perceived Barriers, Cues to Action, Self-Efficacy and Health Behavioural Intention were 0.83, 0.81, 0.86, 0.80, and 0.71, respectively in Phase 1 and 0.87, 0.88, 0.86, 0.89, 0.84, 0.87 and 0.86, respectively in Phase 2 (Table 5).

Table 5: Cronbach's alpha of constructs in Phase 1 and 2

Constructs	Cronbach's Alpha	
	Project 1 (n=100)	Project 2 (n=171)
Perceived Susceptibility	0.83	0.87
Perceived Severity	0.81	0.88
Perceived Benefits	0.86	0.86
Perceived Barriers	0.80	0.89
Cues to Action	0.81	0.84
Self-Efficacy	0.85	0.87
Health Behavioural Intention	0.71	0.86

Discussion

At the time of the research, there are no published papers based on HBM that evaluate cancer health perception among at-risk smokers. Geographic, ethnic groups, national, social, and genetic-related factors contribute to the disproportionate burden of cancer. Sarawak's population vulnerability to NPC is among the highest in the world. Although genetic predisposition may be the most important risk factor leading to higher incidence of NPC in Sarawak, individual behavioural variables are a key driver of community health that should not be underestimated [34].

A total of 34 items in the questionnaire were formulated consistent with the HBM and divided into two sections: HBM scale for smokers' perception to NPC; and health behavioural intention to quit smoking. Both sections are constructed with a 5-point Likert scale ranging from "1 = strongly disagree" to "5 = strongly agree". The first section consists of 30 items and is arranged into six subcategories, each representing the six constructs of the health belief model – perceived susceptibility (5 items), perceived seriousness (5 items), perceived benefits (4 items), perceived barriers (6 items), cues to

action (5 items), and self-efficacy (5 items). The second section includes 4 items that predict the health behavioural intention (See Table S1 and S2).

Based on the validity and reliability tests, including face and content validity, con-struct validity, and internal consistency, the findings of the current study indicate that the questionnaire has shown promising psychometric properties. Ten experts were ad-vised on content validity, and 7 items that did not reach the threshold of CVR based on Lawshe's Table and were judged superfluous were removed [25]. Face validity was examined in a pilot study with 10 smokers who fulfilled the eligibility requirements to ensure cultural acceptance and assess relevance and readability within the local community. Given a satisfactory Impact Score of each item, there was no elimination of item in the face validity stage. In the main study, at-risk smokers from various ethnic groups participated, which was conducted in urban, suburban and rural regions of Miri (the northern region of Sarawak). The KMO test yielded a result of 0.697 (Phase 1) and 0.830 (Phase 2) while the Bartlett's test of sphericity obtained 1746.76 (Phase 1) and 3362.86 (Phase 2), both with p-value < 0.001, indicating that the sample size was adequate and the correlation between the items was sufficient for factor analysis.

Construct validity primarily concerns the degree to which a concept measures what it claims to measure [35]. In parallel to a previous study [36], a number of analyses were conducted to assess the construct validity, including EFA, CFA, as well as convergent and discriminant validity. The EFA demonstrated in Phase 1 that the seven-factor structure accounted for 63.01 percent of the overall variance. The cut-off point for factor loading was fixed at 0.30. According to EFA suggestion, HBI2 item (I am trying to quit smoking to prevent nose and throat cancer at this time) should be grouped together with Perceived Benefit since the factor loading (0.537) is higher than when it is grouped with Health Behavioural Intention (0.307). This item requires immediate smoking cessation, which may spark inconsistent answers as some respondents may not be prepared to quit. The Cronbach's α coefficient value of 'Health Behavioural Intention' construct appeared to be satisfactory (Phase 1: 0.705; Phase 2: 0.861) and this item could potentially be essential in a larger scale. Thus, the researchers agreed not to delete this item.

In Phase 2, CFA was performed to see if the seven-factor model derived by EFA could demystify the association among the items based on the chosen framework. The loadings of all factors were more than 0.5. The goodness of fit was demonstrated to have an acceptable fit with the data with Chi Square = 641.705; df = 500; P< 0.001; CFI = 0.953; TLI: 0.948; RMSEA= 0.041 (90% CI 0.031 to 0.050). Following CFA, convergent and discriminant validity were tested in Phase 2. The CR and AVE values for each component have to be higher than 0.7 and 0.5; respectively, which are fulfilled in our study, suggesting an acceptable convergent validity [37]. Discriminant validity was evaluated between the HBM constructs (excluding health behavioural intention construct). To establish an acceptable discriminant validity, the factors' correlation coefficients with other factors must not be greater than each factor's AVE square root [33]. Our findings demonstrated established discriminant validity for all except for Perceived Bene-fit<-->Cue to Action. The explanation for this might be that the greater the perceived benefits of quitting smoking, the more likely smokers will look for cues to participate in such health-promoting behaviour, or vice versa. Future studies could delve deeper into the strength of the correlation, particularly between Perceived Benefit and Cue to Action. In terms of reliability, each construct for both Phase 1 and 2 showed rationally acceptable Cronbach's α coefficient values as all of which were higher than 0.7, which demonstrate a high internal consistency [38].

This study is not free of limitations. Firstly, it is a cross-sectional study using convenience sampling, and thus, it is susceptible to recall and selection bias. Participants, on the other hand, were given ample time to consider their responses before answering the questions. The second limitation was the relatively small sample size for both EFA (n=100) and CFA (n=171). However, Kline (1994) indicated

that for EFA, a sample size of 100 is sufficient, while Anderson & Gerbing (1988) suggested that CFA/SEM may be reliably examined with a minimum sample size of 100–150 [39, 40]. Finally, the smokers' cultural perspective may be represented in this study. As a result, further studies on smokers from other cultural backgrounds will be required to review the psycho-metric properties of the instrument.

Conclusions

The current study developed a comprehensive HBM-based questionnaire with satisfactory psychometric properties, confirming the validity and reliability. Considering the complex nature of smoking habit, a study with a conceptual framework could comprehensively elucidate the predictive elements and aid us in clearer grasp of this habit, ultimately optimising health promotion [41]. Thus, this questionnaire is developed with the potential of identifying the major obstacles and facilitators to a behaviour change, which can be utilised to inform and design health interventions to im-prove their uptake and efficacy in this population that is at high risk of NPC. With the possibility of being expanded to general health campaigns that target tobacco smoking, the authors also propose further studies to use the instruments for application in other smoking-related cancers in different susceptible populations and geographic locations.

Funding

There was no particular grant for this research from any funding source in the public, private, or non-profit sectors.

Conflicts of Interest

The authors declare that there are no competing or potential conflicts of interest.

Author Contributions

Conceptualization: Kueh Martin, Fairuz Fadzilah Rahim.

Formal analysis: Kueh Martin, Fairuz Fadzilah Rahim.

Investigation: Kueh Martin, Fairuz Fadzilah Rahim.

Methodology: Kueh Martin, Fairuz Fadzilah Rahim.

Project administration: Kueh Martin, Fairuz Fadzilah Rahim.

Resources: Kueh Martin, Fairuz Fadzilah Rahim.

Software: Kueh Martin.

Supervision: Fairuz Fadzilah Rahim.

Validation: Kueh Martin.

Visualization: Kueh Martin.

Writing – original draft: Kueh Martin.

Writing – review & editing: Kueh Martin, Fairuz Fadzilah Rahim, Abdul Rashid.

Institutional Review Board Statement

The study was approved by Joint Penang Independent Ethics Committee of RCSI & UCD Malaysia Campus on 13 August 2020 (Approval no. JPEC 20 0027).

Informed Consent Statement

Informed consent was obtained from all participants involved in the study.

Data Availability Statement

Data supporting the claimed findings can be found at https://datadryad.org/stash/share/QnycI5Un_IWTqoXOr8Zh6xJcsM0RLturj2x7qHAb-ks.

Acknowledgments

We express our gratitude to the hospital director in Miri and local estate managers for their helpful assistance. The participants are sincerely appreciated and much acknowledged.

Word count 3934

References

1. Institute for Health Metrics and Evaluation (IHME). Findings from the Global Burden of Disease Study 2017. In: IHME, editor. Seattle, WA. 2018. Available from: http://www.healthdata.org/policy-report/findings-global-burden-disease-study-2017
2. World Health Organization. WHO Report on Cancer: Setting priorities, investing wisely and providing care for all. 2020. Available from: https://apps.who.int/iris/handle/10665/330745
3. Wellmann KF. [Smoking and Health. On the Report of the Advisory Committee to the Surgeon General of the Public Health Service]. Dtsch Med Wochenschr. 1964, 89, 1085-6.
4. Hecht SS. Research opportunities related to establishing standards for tobacco products under the Family Smoking Prevention and Tobacco Control Act. Nicotine & Tobacco Research. 2011, 14(1), 18-28. https://doi.org/10.1093/ntr/ntq216
5. Viswanath K, Herbst RS, Land SR, Leischow SJ, Shields PG. Tobacco and Cancer: An American Association for Cancer Research Policy Statement. Cancer Research. 2010, 70(9), 3419-30. https://doi.org/10.1158/0008-5472.can-10-1087
6. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021, 71(3), 209-49. https://doi.org/10.3322/caac.21660

7. Hashim D, Boffetta P, La Vecchia C, Rota M, Bertuccio P, Malvezzi M, et al. The global decrease in cancer mortality: trends and disparities. Ann Oncol. 2016, 27(5), 926-33. https://doi.org/10.1093/annonc/mdw027
8. Zavala VA, Bracci PM, Carethers JM, Carvajal-Carmona L, Coggins NB, Cruz-Correa MR, et al. Cancer health disparities in racial/ethnic minorities in the United States. Br J Cancer. 2021, 124(2), 315-32. https://doi.org/10.1038/s41416-020-01038-6
9. Lam KO, Lee AW, Choi CW, Sze HC, Zietman AL, Hopkins KI, et al. Global Pattern of Nasopharyngeal Cancer: Correlation of Outcome With Access to Radiation Therapy. Int J Radiat Oncol Biol Phys. 2016, 94(5), 1106-12. https://doi.org/10.1016/j.ijrobp.2015.11.047
10. Azizah A, Hashimah B, Nirmal K, Siti Zubaidah A, Puteri N, Nabihah A, et al. Malaysian National Cancer Registry Report 2012-2016 [Internet]. Malaysia: National Cancer Institute. . Putrajaya, Malaysia: National Cancer Institute, Ministry of Health, 2019. Available from: https://www.moh.gov.my/moh/resources/Penerbitan/Laporan/Umum/2012-2016%20(MNCRR)/MNCR_2012-2016_FINAL_(PUBLISHED_2019).pdf
11. Devi BCR, Pisani P, Tang TS, Parkin DM. High Incidence of Nasopharyngeal Carcinoma in Native People of Sarawak, Borneo Island. Cancer Epidemiology Biomarkers & Prevention. 2004, 13(3), 482-6.
12. Tsao SW, Yip YL, Tsang CM, Pang PS, Lau VM, Zhang G, et al. Etiological factors of nasopharyngeal carcinoma. Oral Oncol. 2014, 50(5), 330-8. https://doi.org/10.1016/j.oraloncology.2014.02.006
13. Linton RE, Daker M, Khoo ASB, Choo DCY, Viljoen M, Neilsen PM. Nasopharyngeal carcinoma among the Bidayuh of Sarawak, Malaysia: history and risk factors (Review). Oncol Lett. 2021, 22(1), 514. https://dx.doi.org/10.3892%2Fol.2021.12775
14. Guo S-S, Huang P-Y, Chen Q-Y, Liu H, Tang L-Q, Zhang L, et al. The impact of smoking on the clinical outcome of locoregionally advanced nasopharyngeal carcinoma after chemoradiotherapy. Radiation Oncology. 2014, 9(1), 246. https://doi.org/10.1186/s13014-014-0246-y
15. Xue WQ, Qin HD, Ruan HL, Shugart YY, Jia WH. Quantitative association of tobacco smoking with the risk of nasopharyngeal carcinoma: a comprehensive meta-analysis of studies conducted between 1979 and 2011. Am J Epidemiol. 2013, 178(3), 325-38. https://dx.doi.org/10.1093%2Faje%2Fkws479
16. Institute for Public Health. National Health and Morbidity Survey 2015 (NHMS 2015). Ministry of Health Malaysia Kuala Lumpur; 2015. Available from: https://www.moh.gov.my/moh/resources/nhmsreport2015vol2.pdf
17. Rahman M, Arif M, Razak MFA, Suhaili M, Tambi Z, Akoi C, et al. Factors associated with tobacco use among the adult population in Sarawak, Malaysia: A cross sectional study. Epidemiology Biostatistics and Public Health. 2015, 12, 1-9. https://doi.org/10.2427/10292
18. Champion VL, Skinner CS. The health belief model. Health behavior and health education: Theory, research, and practice, 4th ed. San Francisco, CA, US: Jossey-Bass, 2008. p. 45-65.
19. Maiman LA, Becker MH. The Health Belief Model: Origins and Correlates in Psychological Theory. Health Education Monographs. 1974, 2(4), 336-53. https://doi.org/10.1177/109019817400200404
20. Odedina FT, Dagne G, Pressey S, Odedina O, Emanuel F, Scrivens J, et al. Prostate cancer health and cultural beliefs of black men: The Florida Prostate Cancer Disparity Project. Infectious Agents and Cancer. 2011, 6(2), S10. https://dx.doi.org/10.1186%2F1750-9378-6-S2-S10
21. Solomon K, Tamire M, Kaba M. Predictors of cervical cancer screening practice among HIV positive women attending adult anti-retroviral treatment clinics in Bishoftu town, Ethiopia: the application of a health belief model. BMC Cancer. 2019, 19(1), 989. https://doi.org/10.1186/s12885-019-6171-6
22. Tavafian SS, Hasani L, Aghamolaei T, Zare S, Gregory D. Prediction of breast self-examination in a sample of Iranian women: an application of the Health Belief Model. BMC Women's Health. 2009, 9(1), 37. https://doi.org/10.1186/1472-6874-9-37

23. LaMorte., Wayne. W. The Health Belief Model: Behavioral Change Models 2019. Available from: http://sphweb.bumc.bu.edu/otlt/MPH-Modules/SB/BehavioralChangeTheories/BehavioralChangeTheories2.html
24. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine. 2000, 25(24), 3186-91. https://doi.org/10.1097/00007632-200012150-00014
25. Lawshe CH. A quantitative approach to content validity. Personnel psychology. 1975, 28(4), 563-75. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x
26. Samet JM, Wipfli H, Platz EA, Bhavsar N. A Dictionary of Epidemiology, Fifth Edition: Edited by Miquel Porta. American Journal of Epidemiology. 2009, 170(11), 1449-51. https://doi.org/10.1093/aje/kwp322
27. Zamanzadeh V, Ghahramanian A, Rassouli M, Abbaszadeh A, Alavi-Majd H, Nikanfar A-R. Design and Implementation Content Validity Study: Development of an instrument for measuring Patient-Centered Communication. J Caring Sci. 2015, 4(2), 165-78. https://doi.org/10.15171/jcs.2015.017
28. Lacasse Y, Godbout C, Sériès F. Health-related quality of life in obstructive sleep apnoea. Eur Respir J. 2002, 19(3), 499-503. https://doi.org/10.1183/09031936.02.00216902
29. Suhr DD. Exploratory or confirmatory factor analysis? 2006. Available from: https://support.sas.com/resources/papers/proceedings/proceedings/sugi31/200-31.pdf
30. Munro BH. Statistical methods for health care research: lippincott williams & wilkins; 2005.
31. Xiang B, Wong HM, Cao W, Perfecto AP, McGrath CPJ. Development and validation of the Oral health behavior questionnaire for adolescents based on the health belief model (OHBQAHBM). BMC Public Health. 2020, 20(1), 701. https://doi.org/10.1186/s12889-020-08851-x
32. Brown TA. Confirmatory factor analysis for applied research. New York, NY, US: The Guilford Press; 2006.
33. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. Journal of marketing research. 1981, 18(1), 39-50. https://doi.org/10.2307/3151312
34. McGinnis JM, Williams-Russo P, Knickman JR. The case for more active policy attention to health promotion. Health affairs. 2002, 21(2), 78-93. https://doi.org/10.1377/hlthaff.21.2.78
35. Cohen RJ, Swerdlik ME, Phillips SM. Psychological testing and assessment: An introduction to tests and measurement: Mayfield Publishing Co; 1996.
36. Shamsalinia A, Moradi M, Farahani MA, Masoudi R, Ghadimi R, Rad RE, et al. Designing and psychometric evaluation of disease-related fear scale (D-RFS) in adults with epilepsy: A sequential exploratory mixed methods design. Epilepsy & Behavior. 2020, 110, 107169. https://doi.org/10.1016/j.yebeh.2020.107169
37. Schreiber JB, Nora A, Stage FK, Barlow EA, King J. Reporting structural equation modeling and confirmatory factor analysis results: A review. The Journal of educational research. 2006, 99(6), 323-38. https://doi.org/10.3200/JOER.99.6.323-338
38. Kline P. The handbook of psychological testing: Psychology Press, 2000. https://doi.org/10.4324/9781315812274
39. Kline P. An Easy Guide to Factor Analysis (1st ed.). 1st edition ed. London. 1994, p. 208. https://doi.org/10.4324/9781315788135
40. Anderson JC, Gerbing DW. Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. Psychological Bulletin. 1988, 103, 411-423. http://dx.doi.org/10.1037/0033-2909.103.3.411
41. Reisi M, Javadzade S, Shahnazi H, Sharifirad G, Charkazi A, Moodi M. Factors affecting cigarette smoking based on health-belief model structures in pre-university students in Isfahan, Iran. Journal of Education and Health Promotion. 2014, 3(1), 23. https://doi.org/10.4103/2277-9531.127614

Figure

Figure 1. Conceptual framework of the study based on Health Belief Model

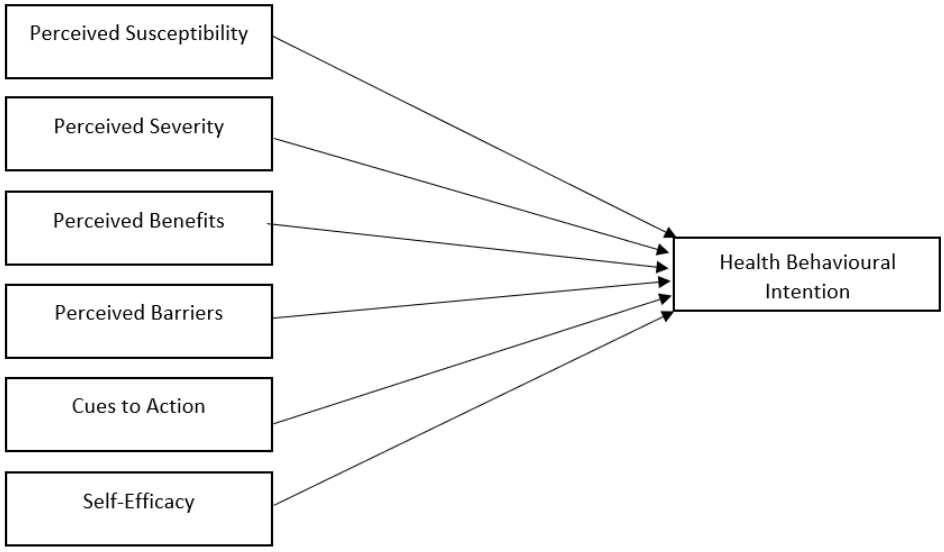

Figure 1. Conceptual framework of the study based on Health Belief Model

144x82mm (157 x 157 DPI)

Table S1: Questionnaire to predict health behavioural intention to prevent nasopharyngeal cancer among at-risk smoker (English Version)

Section 1: Perceptions towards nose and throat cancer

Please choose the best answer options to indicate your level of agreement.

Constructs of Health Belief Model	Questions	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Perceived susceptibility	1. I feel that I will get nose and throat cancer in the future.					
	2. My chances of getting nose and throat cancer are high.					
	3. I cannot avoid myself from getting nose and throat.					
	4. I am worried about getting nose and throat.					
	5. My smoking habit makes me more likely than average to get nose and throat cancer.					
Perceived seriousness	6. The thought of nose and throat cancer scares me.					
	7. If I had nose and throat cancer, the cost of treatment can be a financial strain.					
	8. Nose and throat cancer would threaten a relationship with my family.					
	9. If I had nose and throat cancer my					

	whole life would change.					
	10. If I developed nose and throat cancer, I would not live long.					
Perceived benefits	11. Quit smoking decreases my chance of getting nose and throat cancer.					
	12. Quit smoking decreases my chance of dying from nose and throat cancer.					
	13. Quit smoking can improve my health.					
	14. I feel less anxious about nose and throat cancer if I quit smoking.					
Perceived barriers	15. It is difficult to quit tobacco smoking (e.g. peer pressure).					
	16. I feel anxious without smoking,					
	17. Tobacco smoking relieves my stress.					
	18. I experience headache without smoking.					
	19. I experience excessive salivation without smoking.					
	20. I think getting nose and throat cancer is destiny and quitting smoking will not change it.					
Cues to action	21. I will stop me from smoking if I have social support.					
	22. I will stop smoking if there are					

	information sources that reminds me. Examples of sources include: the internet, newspapers, radio and TV.					
	23. I will stop smoking if I have the will to change.					
	24. I will stop smoking if I know the diseases related to smoking.					
	25. I will stop smoking if there are health professionals to assist me.					
Self-efficacy	26. I can refuse to smoke when I am thinking about difficult problem.					
	27. I can refuse the urge to smoke.					
	28. I can refuse to smoke when I see someone else smoking.					
	29. I can refuse to smoke when offered by my friends/ family.					
	30. I can refuse from buying cigarettes when I have extra pocket money.					

Section 2: Health Behavioural Intention to quit smoking

Please choose the best answer options to indicate your level of agreement.

Statements	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
1. I would like to lead a healthier lifestyle to prevent nose and throat cancer.					

2. I am trying to quit smoking to prevent nose and throat cancer at this time.					
3. I plan to quit smoking to prevent nose and throat cancer within six months.					
4. I would like to quit smoking to prevent nose and throat cancer but have never really tried.					

Table S2: Questionnaire to predict health behavioural intention to prevent nasopharyngeal cancer among at-risk smoker (Malay Version)

Bahagian 1: Persepsi terhadap kanser hidung dan tekak.

Sila bulatkan dengan pilihan jawapan terbaik untuk menunjukkan tahap persetujuan anda.

Domain Model Kepercayaan Kesihatan	Soalan	Sangat tidak setuju	Tidak setuju	Kurang bersetuju	Setuju	Sangat setuju
Kepercayaan kepada tiada daya tahan	1. Saya merasakan bahawa saya akan mendapat kanser pangkal hidung pada masa akan datang.					
	2. Peluang saya mendapat kanser pangkal hidung adalah tinggi.					
	3. Saya tidak dapat mengelakkan diri daripada mendapat kanser pangkal hidung.					
	4. Saya bimbang terkena kanser pangkal hidung kerana ahli keluarga saya menghadapi masalah ini.					
	5. Tabiat merokok saya menyebabkan saya lebih berisiko menghidap kanser pangkal hidung.					
Kepercayaan kepada bahaya	6. Risiko dan kesan-kesan kanser pangkal hidung menakutkan saya.					
	7. Jika saya mempunyai kanser pangkal hidung, kos rawatan akan					

	menjadi beban kewangan keluarga saya.					
	8. Kanser pangkal hidung akan merosakkan hubungan antara ahli keluarga saya.					
	9. Jika saya mempunyai kanser pangkal hidung, kehidupan saya akan diubah.					
	10. Jika saya mempunyai kanser pangkal hidung, saya tidak akan hidup lama.					
Kepercayaan kepada manfaat	11. Berhenti merokok akan mengurangkan peluang saya mendapat kanser pangkal hidung.					
	12. Berhenti merokok akan mengurangkan risiko kematian disebabkan kanser pangkal hidung.					
	13. Berhenti merokok amat penting untuk meningkatkan kesihatan saya.					
	14. Berhenti merokok akan mengurangkan kerisauan saya terhadap kanser pangkal hidung.					
Kepercayaan kepada halangan	15. Saya berasa sukar untuk menghentikan tabiat merokok atas sebab tekanan					

	rakan sebaya dan lain-lain.					
	16. Saya akan berasa bimbang akan sesuatu tanpa merokok.					
	17. Merokok menghilangkan tekanan saya.					
	18. Jika saya tidak merokok, saya akan mengalami sakit kepala					
	19. Jika saya tidak merokok, saya akan mengalami air liur berlebihan.					
	20. Saya rasa mendapat kanser pangkal hidung adalah takdir dan berhenti merokok tidak akan mengubahnya.					
Isyarat untuk bertindak	21. Jika saya mempunyai sokongan sosial, saya akan berhenti merokok.					
	22. Jika beberapa sumber maklumat mengingatkan saya, saya akan berhenti merokok. Contoh sumber maklumat termasuk internet, surat khabar, radio dan TV.					
	23. Saya akan berhenti merokok sekiranya saya mempunyai kemahuan untuk berubah.					
	24. Menyedari bahaya merokok					

	akan membantu saya berhenti merokok.					
	25. Saya akan berhenti merokok sekiranya ahli profesional kesihatan membantu saya					
Keberkesanan diri	26. Saya boleh menolak merokok apabila saya ingin memikirkan masalah yang sukar.					
	27. Saya boleh menolak merokok apabila saya mempunyai keinginan untuk merokok.					
	28. Saya boleh menolak merokok apabila saya melihat orang lain merokok.					
	29. Saya boleh menolak merokok apabila kawan/ keluarga saya mengajak saya untuk merokok.					
	30. Saya boleh menolak membeli rokok apabila saya mempunyai lebihan wang saku.					

Bahagian 2: Tingkah laku untuk berhenti merokok demi kesihatan

Sila bulatkan dengan pilihan jawapan terbaik untuk menunjukkan tahap persetujuan anda.

Soalan	Sangat tidak bersetuju	Tidak setuju	Kurang bersetuju	Setuju	Sangat setuju
1. Saya akan menjalani gaya hidup yang lebih sihat untuk mencegah kanser pangkal hidung.					

2. Saya sedang berusaha untuk berhenti merokok untuk mengurangkan risiko kanser pangkal hidung pada masa ini.					
3. Saya merancang untuk berhenti merokok untuk mengurangkan risiko kanser pangkal hidung dalam masa enam bulan.					
4. Saya ingin berhenti merokok untuk mengurangkan risiko kanser pangkal hidung tetapi tidak pernah mencuba.					

Reporting checklist for cross sectional study.

Based on the STROBE cross sectional guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the STROBE cross sectionalreporting guidelines, and cite them as:

von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies.

		Reporting Item	Page Number
Title and abstract			
Title	#1a	Indicate the study's design with a commonly used term in the title or the abstract	1
Abstract	#1b	Provide in the abstract an informative and balanced summary of what was done and what was found	1
Introduction			
Background / rationale	#2	Explain the scientific background and rationale for the investigation being reported	2-3
Objectives	#3	State specific objectives, including any prespecified hypotheses	3
Methods			
Study design	#4	Present key elements of study design early in the paper	5

Setting	#5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Eligibility criteria	#6a	Give the eligibility criteria, and the sources and methods of selection of participants.	5
	#7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	4-5
Data sources / measurement	#8	For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Give information separately for for exposed and unexposed groups if applicable.	5-6
Bias	#9	Describe any efforts to address potential sources of bias	5 - 6. We involved experts and participants of different background and ethnic groups in hopes to reduce biasness.
Study size	#10	Explain how the study size was arrived at	12
Quantitative variables	#11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	6
Statistical methods	#12a	Describe all statistical methods, including those used to control for confounding	6
Statistical methods	#12b	Describe any methods used to examine subgroups and interactions	6
Statistical methods	#12c	Explain how missing data were addressed	N/A. No missing data was collected.
Statistical methods	#12d	If applicable, describe analytical methods taking account of sampling strategy	6
Statistical methods	#12e	Describe any sensitivity analyses	N/A

Results

Participants	#13a	Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed. Give information separately for for exposed and unexposed groups if applicable.	N/A. Do not have a list of numbers potentially eligible.
Participants	#13b	Give reasons for non-participation at each stage	N/A
Participants	#13c	Consider use of a flow diagram	N/A. Do not have a list of numbers potentially eligible.
Descriptive data	#14a	Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Give information separately for exposed and unexposed groups if applicable.	7
Descriptive data	#14b	Indicate number of participants with missing data for each variable of interest	N/A. No missing data was collected
Outcome data	#15	Report numbers of outcome events or summary measures. Give information separately for exposed and unexposed groups if applicable.	N/A. This study focuses on validating the questionnaire with a series of analysis.
Main results	#16a	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	N/A. This study focuses on validating the questionnaire with a series of analysis.
Main results	#16b	Report category boundaries when continuous variables were categorized	N/A. This study focuses on validating the questionnaire with a series of analysis.
Main results	#16c	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A. This study focuses on validating the questionnaire with a series of analysis.
Other analyses	#17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	8-11

Discussion			
Key results	#18	Summarise key results with reference to study objectives	11-12
Limitations	#19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.	12
Interpretation	#20	Give a cautious overall interpretation considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.	12
Generalisability	#21	Discuss the generalisability (external validity) of the study results	12
Other Information			
Funding	#22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	13

Notes:

- 9: 5 - 6. We involved experts and participants of different background and ethnic groups in hopes to reduce biasness.
- 12c: N/A. No missing data was collected.
- 13a: N/A. Do not have a list of numbers potentially eligible.
- 13c: N/A. Do not have a list of numbers potentially eligible.
- 14b: N/A. No missing data was collected
- 15: N/A. This study focuses on validating the questionnaire with a series of analysis.
- 16a: N/A. This study focuses on validating the questionnaire with a series of analysis.
- 16b: N/A. This study focuses on validating the questionnaire with a series of analysis.
- 16c: N/A. This study focuses on validating the questionnaire with a series of analysis. The STROBE checklist is distributed under the terms of the Creative Commons Attribution License CC-BY. This

checklist was completed on 19. September 2021 using https://www.goodreports.org/, a tool made by the EQUATOR Network in collaboration with Penelope.ai

BMJ Open

Development and validation of the health behavioural intention on smoking cessation to prevent nasopharyngeal cancer in Sarawak, Malaysia based on the Health Belief Model: a cross-sectional study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-057552.R1
Article Type:	Original research
Date Submitted by the Author:	20-May-2022
Complete List of Authors:	Kueh, Martin; RCSI & UCD Malaysia Campus, Department of Public Health Medicine Rahim, Fairuz Fadzilah; RCSI & UCD Malaysia Campus, Department of Public Health Medicine Rashid, Abdul; RCSI & UCD Malaysia Campus, Public Health Medicine
Primary Subject Heading:	Public health
Secondary Subject Heading:	Public health, Smoking and tobacco, Oncology
Keywords:	PUBLIC HEALTH, Head & neck tumours < ONCOLOGY, Epidemiology < ONCOLOGY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Development and validation of the health behavioural intention on smoking cessation to prevent nasopharyngeal cancer in Sarawak, Malaysia based on the Health Belief Model: a cross-sectional study

Martin Kueh[1], Fairuz Fadzilah Rahim[1] & Abdul Rashid[1]

[1] Department of Public Health Medicine, RCSI & UCD Malaysia Campus, No.4, Jalan Sepoy Lines, 10450 George Town, Penang Malaysia

Correspondence: Martin Kueh, Department of Public Health Medicine, RCSI & UCD Malaysia Campus, No.4, Jalan Sepoy Lines, 10450 George Town, Penang Malaysia. Tel: 6016-879 2269. Email: martin1kueh@hotmail.com

Abstract

Objective Lifestyle-induced nasopharyngeal carcinoma (NPC) is a serious but preventable risk factor. This study serves to develop and validate a questionnaire that aims to predict the health behavioural intention on smoking cessation in Sarawak, Malaysia using the Health Belief Model (HBM).

Design A cross-sectional study

Setting Urban and suburban areas in Sarawak, Malaysia.

Participants The preliminary items of the instrument were developed through a literature review. The instrument was translated into Malay language using forward-backwards method before conducting the content validity through a panel of 10 experts. Face validity was examined both quantitatively and qualitatively by 10 local smokers. The construct validity of the instrument was evaluated through exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). A total of 100 local smokers participated in Phase 1 for EFA, while 171 local smokers participated in Phase 2 for CFA. Internal consistency was measured using Cronbach's alpha coefficients to evaluate the reliability.

Results In the exploratory stage, the factor loading of each item remained within the acceptable threshold. The final revised CFA yielded appropriate fit of the seven-factor model with the following model fit indexes: Chi-Square: 641.705; df= 500; P< 0.001; CFI = 0.953; TLI: 0.948; RMSEA= 0.041. Satisfactory convergent validity and divergent validity were shown, with the exception of one pairwise construct. The internal reliability of these scales was above the desirable threshold, with Cronbach's alpha coefficients ranging from 0.705 to 0.864 and 0.838 to 0.889 for Phase 1 and 2, respectively.

Conclusions The study substantiated the instrument to be valid and reliable for predicting smokers' health behavioural intention to reduce cancer risk. The instrument is made up of 34 items, categorised into two sections, six HBM constructs and health behavioural intention. The instrument can be utilised for other smoking cessation-related cancers in different at-risk population.

Keywords: tobacco smoking, cancer health promotion, nasopharyngeal cancer, health belief model, development and validation, quit smoking, smoking cessation

Article Summary

Strengths and limitations of this study

1. This study established a novel instrument for assessing at-risk smokers' health behavioural intention to reduce NPC risk based on a well-known framework in a series of systematic validation stages, which potentially can be applied for other smoking-related cancers.

2. Face validity was undertaken both qualitatively and quantitatively to sufficiently reflect the demographic during the assessments of psychological constructs. The validation was markedly aided by experts' evaluations.

3. This study was conducted in two phases, involving both urban and suburban smokers, to examine concept validity, convergent validity, and divergent validity.

4. The study's generalizability may be limited by the smokers' cultural perspective, hence, further studies on smokers from different cultures will be needed to assess the instrument's psychometric properties.

Introduction

According to the World Health Organization, tobacco smoking is a public health concern that accounts for over 8 million deaths per year and is the leading avoidable cause of illness, disability, and death globally [1]. Annually, exposure to smoking is associated with 2.4 million deaths from cancer throughout the world [2]. The report from the Surgeon General of the United States of America (USA) associating smoking and cancer was a watershed moment in public health towards tobacco's adverse effects on human health, followed by subsequent discovery that tobacco smoke comprises approximately 7,000 compounds, 72 of which are carcinogenic [3, 4]. Tobacco use is now causally associated with at least 20 cancer types, which also appears to have a wide-ranging immediate and long-term health benefits accompanying smoking cessation [2]. However, the harmful consequences of tobacco smoking are widely neglected or underestimated, despite the fact that it remains a significant public health hazard among with the impoverished, and marginalised, as well as those in developing nations, which bear a disproportionate share of the burden [5].

Cancer is a leading cause of death as well as a major obstacle to improving lifespan in every country [6]. There is an approximately 1% reduction in the overall cancer mortality rate across both sexes in both high- and low-income countries [7]. However, this is mainly a result of positive trends in the most common cancers. It fails to account for variations in the frequency and distribution of aetiologic aspects such as socioeconomic, geographical, genetic, biological, ethnic, social, and physical factors, as well as disparities across cancer types [7, 8]. For instance, nasopharyngeal cancer (NPC) is uncommon but unique among head and neck cancers and is a cancer disparity with its own distinct epidemiological & risk factor profiles. To reinforce the claim, the Global data from the World Health Organization illustrates poorer outcomes of NPC in endemic areas like Southeast Asia with an unbalanced global burden of 67% [9].

In Malaysia, NPC is a nationwide public health concern and the 5th leading form of cancer, amounting to 4,597 new diagnoses of NPC for the 2012-2016 period. The recent report from the Malaysian National Cancer Registry reported that the lifetime risk of developing NPC among men and women are 1 in 175 and 1 in 482, respectively [10]. Strikingly, there is a substantial geographical variance within the country, with Sarawak exhibiting a higher prevalence rate of NPC. A previous study has

shown a significant high age-standardised rates in males (13.5/100,000, 95% Confidence Interval = 12.2 – 15.0) and females (6.2/100,000, 95% CI= 5.7-6.7) by which the local at-risk ethnic groups including Bidayuh, Chinese, Iban, Malays and Melanau were collectively ranked top globally. In particular, the risk among the Bidayuh ethnic population, which is a native indigenous group, exceeds the general population of males and females in Sarawak by 2.3 times and 1.9 times respectively [11].

This trend has been ascribed to potential risk factors, which include Epstein-Barr virus, genetic susceptibility, consumption of food with nitrous compounds and volatile nitrosamines, and complex interaction with environmental factors [12]. Among many risk factors that are associated with NPC, tobacco smoking is the most important modifiable cause of the severity of NPC [13, 14]. A meta-analysis 32 epidemiologic studies (28 case-control studies and 4 cohort studies) regarding tobacco smoking and NPC from 1979 to 2011 reported that tobacco-correlated NPC cases were 60% higher compared to non-smokers [15]. The Malaysian National Health and Morbidity Survey 2015 showed that the prevalence of tobacco smoking among the population in Sarawak was 25.4%. The native indigenous male (61.2%) and female smokers (10.7%) in Sarawak were among the highest nationwide [16]. Rahman et al (2015) indicated that the average number of tobaccos smoked was 13.6 cigarettes per day in Sarawak [17]. This lifestyle-induced NPC is a serious concern, and necessitates a preventative strategy centred on modifying health risk behaviours.

Despite the robust establishment of cumulative impact of tobacco smoking on the risk of cancer, there is still a paucity of local studies and research efforts to discover critical areas and serve as a benchmark for the evaluation of the effectiveness of comprehensive strategies to promote cancer prevention among smokers. Hence, given the significance of this subject, it is imperative to create a questionnaire that focuses on behavioural factors and is customised to the interests of local at-risk smokers. The Health Belief Model (HBM), an approachable theoretical model that could aid in the understanding of the individual's or a smoker's belief on the health-related behavioural intention [18, 19]. This HBM can further predict the smokers' effort to improve health or their health-seeking behaviours in the preventive practices of NPC. This study serves to develop and validate a health behaviour model (HBM)-based questionnaire that aims to predict the health behavioural intention of these smokers and their perspective and motivation towards smoking cessation.

Conceptual Framework

The Health Belief Model (HBM) is the underpinning conceptual framework for this study to focus on psychological variables for predicting health behavioural intention. Developed in the 1950s by social psychologists at the United States Public Health Service, the Health Belief Model (HBM) is currently one of the most extensively used cognitive model and theoretical framework to help researchers understand and predict health behaviours in the population and ultimate guide for health promotion and interventions activities [18, 19]. A large volume of studies in the past utilized HBM to examine health-promoting behaviours for the prevention of different cancers in both developed and developing countries including, the USA, Iran and Ethiopia [20-22]. The HBM is a value-expectancy theory, based on the hypothesis of Lewin et al that highlights the influence of two variables on behaviour: 1) the value that a person places on the outcome of the behaviour and 2) the person's perception of how likely the behaviour will lead to that outcome, in the event of an illness (19). Having evolved over the past decades, HBM currently consists of 6 elements: i) perceived susceptibility, ii) perceived severity, iii) perceived benefits, iv) perceived barriers, v) cues to action, and vi) self-efficacy [23].

The first four refer to a person's subjective perceptions regarding 1) his/her risk of getting the disease; 2) how severe the consequences are of getting the disease; 3) the benefits from performing a health behaviour in preventing, curing, or managing the disease; and 4) obstacles to that health behaviour, e.g. financial and time costs, side effects, and so on (23). "Cue to action" is the stimulus, which may be internal (e.g., physical sensations) or external (e.g., friends with the disease, social media), that is required for that health behaviour to occur, and "self-efficacy" refers to the person's confidence on how capable he/she is to successfully undertake that health behaviour [19, 23].

Methods

Study design and setting

A cross-sectional study was conducted in Sarawak, Malaysia in two phases: Phase 1 from October 2020 to January 2021 and Phase 2 from January to April 2021. Sarawak is the largest Malaysian state on the island of Borneo with a population of more than 2.6 million, made up of 26 different ethnic groups. Sarawak is divided into twelve divisions, each of which is further divided into districts and sub-districts.

The sample population of Phase 1 were residents residing in urban and suburban areas in Miri, Sarawak. Phase 2 mainly involved by residents residing in Bintulu, Kuala Baram (a federal constituency in Miri Division), and remote rural areas in Miri. They were mainly local employees working in the agricultural industry. Data were collected by eight trained research assistants. All research assistants were given a crash course in the research aims, methodology and data collection, as well as a trial run to simulate real-world situations. Prior to the distribution of questionnaires, participants were briefed regarding the objective and methodology of the study, as well as the benefits and risks involved. Involvement in this sample was entirely voluntary and did not pose any potential threat. Upon clarifying the study details, informed consent was obtained from each participant. The anonymity of the respondents' details is assured. The research assistants provided clarifications to smokers who requested assistance. The methodology of the study and data collection were recorded precisely and accurately throughout the process of the research.

The inclusion criteria were: 1) 18 years old and above, 2) smoked for at least a year, and 3) is a Sarawakian. Participants who did not consent to participate, were pregnant, or smoked e-cigarettes only were excluded from this study.

Patient and Public Involvement

No patient was involved in this study.

Instrument development and validation

The questionnaire was self-administered based on prior validated studies. Search was conducted using the National Library of Medicine (PubMed), Google Scholar and Cochrane Library databases by exploring various keywords: health belief model, nasopharyngeal cancer, cancer, smokers, smoking behaviours/habit, and questionnaire/tools/instruments. Questionnaires and prior literatures were examined primarily, and explicitly on smokers around about NPC prevention using HBM. The initial questionnaire consisted of 41 items.

The questionnaire was developed in English and translated to Malay by two local fluent bilingual translators. An experienced researcher, whose mother tongue is Malay, compared the Malay version of the questionnaire to the English version. A 'back-translation' approach to English was taken independently by another two bilingual translators based on the Beaton-recommended guidelines [24].

To determine the content validity, 10 healthcare professionals including public health experts, hospital directors, and health officers were invited to evaluate the survey instrument. According to Lawshe's model, a questionnaire was designed and organized to assist and allow the panellists to express clearly

their views on the importance of including different components in a model. Experts received attachments of the questionnaire via email, which was graded on a three-point scale: essential, useful but not essential, and not essential. Based on Lawshe's table, items with a CVR value greater than 0.62 were retained [25]. To minimise ambiguity, the experts evaluated each item's accuracy, phrasing and grammar as well as their relevance to the construct. Modifications were made for subsequent analysis based on the experts' comments and suggestions. Overall, 7 items were deleted, 2 items were rephrased and 1 item was allocated into a different construct.

Face validity was evaluated through a pilot study by 10 local smokers of different ethnic background, both qualitatively and quantitatively. Smokers who participated in the pilot study were exempted from the main study. The pilot study is conducted on a small-scale basis to ascertain the feasibility of the proposed larger study [26]. In the qualitative stage, cognitive interviews were conducted face-to-face individually to obtain participant's feedback on their comprehension and answers. Items which were not well understood were identified from the cognitive interview [27]. Minor revisions were made to better suit a linguistically and culturally diverse context. Based on the local smokers' perspective, an item-oriented towards a person's faith or spirituality to own health was also included. With the consensus of the researchers, an item was added, 'I think getting nose and throat cancer is my destiny and quitting smoking will not change it'. In the quantitative stage, a survey was disseminated based on a Likert scale of 1 (least importance) to 5 (extremely important) to determine the clinical impact of each item. The importance score was calculated based on "clinical impact method" in which the clinical impact of each item was determined from the proportion of participants who identified it as important. This technique was chosen for better clarity where the items were ranked according to their impact score. Mean importance score of each item was computed using the following formula: Impact Score = Frequency (Proportion) x Importance. Factors were kept if the Impact Score equal or more than 1.5. They were defined as deemed suitable and kept for further evaluation [28]. In the current study, the impact score for each item ranged from 1.7 to 4.6, therefore, no item was eliminated.

Subsequently, in the main study, to determine the construct validity, exploratory and confirmatory factor analyses were performed in two periods. In the first period, EFA (n = 100) was used to determine the number of latent factors or the relationships between the common factors. The model was later adjusted in the second period with CFA (n = 171) via structural equations using AMOS. CFA confirmed the overall fit of the model and indicated that the measures were in the acceptable range [29]. Convergent validity and discriminant validity were also carried out in Phase 2.

Final scoring of the instrument:

In the main study, the total mean score for each HBM components were formulated on the five-point Likert scale options from strongly disagree to strongly agree, with scores ranging from 1 to 5 points. With the exception of 'Perceived Barriers,', which is inversely proportional, a greater score reflects a firmer desire to quit smoking.

Ethical considerations:

This study was approved by RCSI & UCD Malaysia Campus (RUMC) Institutional Research and Ethics Committee (Approval no. JPEC 20 0027). The informed consent of all participants was obtained voluntarily, and the data confidentiality and storage are assured.

Data management:

The data was processed in Microsoft Excel before being analysed in SPSS 26.0 and AMOS 23.0 [30]. Listwise deletion approach was done for missing data of less than 2%. Socio-demographic

characteristics are presented as number and percentage distribution. Cronbach's alpha coefficients was used to determine internal consistency. Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) used to test the construct validity of each construct. Specifically, EFA was evaluated in Phase 1 and CFA in Phase 2. A p-value below 0.05 was deemed to be statistically significant.

EFA was performed in Phase 1 to reveal the fundamental structure of a large set of variables [29]. The factors were extracted using principal component analysis with a varimax rotation. As for sampling adequacy and item-checking, the Kaiser-Meyer-Olkin (KMO) > 0.6 and Bartlett's test for sphericity ($p < 0.05$) were used. All loading factors below 0.3 were excluded from the constructs [31].

CFA was performed in Phase 2 to assess the data integrity and the structural model [29]. The acceptable level of standardised factor loading was set at 0.5 and above to ensure a satisfactory association between items and corresponding factors [32]. Different fit indices were utilised to estimate the model fit, for instance a comparative fit index (CFI) of > 0.90, Tucker-Lewis Index (TLI) > 0.90 and root mean square error of approximation (RMSEA) < 0.05 [33]. Additionally, convergent validity and discriminant validity were evaluated based on the composite reliability (CR) and average variance extracted (AVE). The HBM components were analysed with the minimum and maximum scores, total mean score and standard deviation (SD).

Results:

A total of 100 and 171 smokers participated in Phase 1 and Phase 2 of the study, respectively. The response rate for Phase 1 was 100% (100/100), whereas Phase 2 response rate was 98.3% (171/174). Majority of the participants were males (Phase 1 86.0%; Phase 2 89.5%) and were in the 30-39 age group (Phase 1 40.0%; Phase 2 40.9%). A little over a third of the participants smoked for more than 10 years for both Phase 1 (37%) and Phase 2 (40.4%). In Phase 1, most of the participants were Iban (34%), followed by Malay (19%), Chinese (18%), Others (13%), Melanau (9%) and Bidayuh (7%). Most of the participants in Phase 2 were Chinese (33.9%), followed by Iban (24.6%), Malay (12.3%), Bidayuh (9.9%), Melanau (9.9%) and Others (9.4%). Details of the smokers' demographic information are presented in Table 1.

Table 1: Demographic characteristics of participants

	Phase 1 (n=100)		Phase 2 (n=171)	
	Number	Percentage (%)	Number	Percentage (%)
Age				
18-29	34	34.0	46	26.9
30-39	40	40.0	70	40.9
40-49	14	14.0	43	25.1
50-64	12	12.0	10	5.8
65 and above	0	0.0	2	1.2
Gender				
Male	86	86.0	153	89.5
Female	14	14.0	18	10.5
Ethnic groups				
Malay	19	19.0	21	12.3
Chinese	18	18.0	58	33.9
Bidayuh	7	7.0	17	9.9

Iban	34	34.0	42	24.6
Melanau	9	9.0	17	9.9
Others	13	13.0	16	9.4
Years of smoking				
1-5 years	32	32.0	44	25.7
6-10 years	31	31.0	58	33.9
More than 10 years	37	37.0	69	40.4

In Phase 1, Kaiser-Meyer-Olkin Measure of Sampling Adequacy was 0.697 and Barlett's test of the sphericity was significant (x2 = 1,746, p-value < 0.001). EFA was conducted to analyse the factor structure with principal component analysis with a varimax rotation. A decision was made to go for 7-factor structures since there is clarity of 7 constructs. The EFA found that 7 variables had eigenvalues larger than Kaiser's threshold of 1 and explained 63.0% of the variance when combined. Factor loadings of higher than 0.3 were found in all the items. Four items had cross loading with values greater than 0.3, which are PBar5, HBI2, HBI3 & HBI4. All items remained because the contents of the items were regarded as relevant based on the decision and judgment of the researchers. Table 2 shows the EFA with total items and the factor loading of each construct for the 7-factor model.

Table 2: Result of Exploratory Factor Analysis in Phase 1 (n = 100)

		Component						
Constructs	Items	1	2	3	4	5	6	7
Perceived Susceptibility	PSus1			0.665				
	PSus2			0.750				
	PSus3			0.810				
	PSus4			0.548				
	PSus5			0.736				
Perceived Severity	PSev1						0.558	
	PSev2						0.720	
	PSev3						0.744	
	PSev4						0.843	
	PSev5						0.730	
Perceived Benefit	PBen1	0.767						
	PBen2	0.765						
	PBen3	0.763						
	PBen4	0.792						
Perceived Barrier	PBar1					0.594		
	PBar2					0.725		
	PBar3					0.658		
	PBar4					0.709		
	PBar5					0.590		
	PBar6					0.617		
Cue to action	CUE1				0.721			
	CUE2				0.584			
	CUE3				0.736			
	CUE4				0.723			
	CUE5				0.745			

Self-efficacy	EFF1		0.721					
	EFF2		0.730					
	EFF3		0.770					
	EFF4		0.868					
	EFF5		0.625					
Health Behavioural Intention	HBI1							0.681
	HBI2							0.307
	HBI3							0.733
	HBI4							0.596
Rotation Sums of Squared Loading								
Total		3.469	3.371	3.246	3.072	3.071	3.012	2.185
Percentage of Variance		10.203	9.913	9.548	9.034	9.032	8.859	6.425
Cumulative percentage		10.203	20.116	29.664	38.698	47.730	56.588	63.014

CFA was performed in Phase 2 to assess whether the seven-factor model indicated by the EFA could sufficiently represent the data. The items in their respective constructs were loaded between 0.586 and 0.898 (Table 3). For the model's fitness to increase, items with less than 0.6 and a MI more than 10 should have been eliminated. Despite this, they were kept because they were essential for the conceptual framework. Before arriving at the final model, 6 pairs of correlated errors were added to improve robustness. The resulting model was suitable for testing, as evidenced by the following model fit indexes: Chi Square: 641.705; df= 500; P< 0.001; CFI = 0.953; TLI: 0.948; RMSEA= 0.041 (90% CI 0.031 to 0.050).

Table 3: Result of Confirmatory Factor Analysis in Phase 2 (n = 171)

Constructs	Items	Factor Loadings	AVE	CR
Perceived Susceptibility	PSus1	0.898	0.577	0.871
	PSus2	0.819		
	PSus3	0.739		
	PSus4	0.685		
	PSus5	0.683		
Perceived Severity	PSev1	0.680	0.597	0.881
	PSev2	0.675		
	PSev3	0.806		
	PSev4	0.867		
	PSev5	0.766		

Perceived Benefit	PBen1	0.752	0.603	0.858
	PBen2	0.812		
	PBen3	0.740		
	PBen4	0.733		
Perceived Barrier	PBar1	0.586	0.572	0.888
	PBar2	0.670		
	PBar3	0.626		
	PBar4	0.886		
	PBar5	0.878		
	PBar6	0.770		
Cue to action	CUE1	0.646	0.512	0.839
	CUE2	0.721		
	CUE3	0.727		
	CUE4	0.767		
	CUE5	0.723		
Self-efficacy	EFF1	0.645	0.574	0.869
	EFF2	0.694		
	EFF3	0.827		
	EFF4	0.859		
	EFF5	0.744		
Health Behavioural Intention	HBI1	0.654	0.617	0.864
	HBI2	0.836		
	HBI3	0.867		
	HBI4	0.773		

The AVE and CR values, which are listed in Table 3, were obtained after the structural model's fit was investigated to check if the items were loaded appropriately. The AVE readings were all over the cut-off value of 0.5, ranging from 0.512 to 0.617. The CR values were all over the cut-off value of 0.7, ranging from 0.839 to 0.888. All seven constructs featured sufficient convergent validity.

Discriminant validity was evaluated using Fornell-Larcker criteria by comparing the squared correlations and AVE scores for each of the pairwise constructs [34]. With the exception of Perceived Benefit < — > Cue to action, all paired constructs have shown established discriminant validity (Table 4).

Table 4: Result of Discriminant Validity in Phase 2 (n = 171)

	Factor Correlation	Correlation Squared	Discriminant validity

Perceived Susceptibility<-->Perceived Severity	.312	0.097	Established
Perceived Susceptibility <-->Perceived Benefit	.260	0.068	Established
Perceived Susceptibility <-->Perceived Barrier	.204	0.042	Established
Perceived Susceptibility <-->Cue to Action	.172	0.030	Established
Perceived Susceptibility <-->Self-Efficacy	.236	0.056	Established
Perceived Severity <--> Perceived Benefit	.575	0.331	Established
Perceived Severity <--> Perceived Barrier	-.042	0.176	Established
Perceived Severity <--> Cue to Action	.238	0.057	Established
Perceived Severity <--> Self-Efficacy	.257	0.066	Established
Perceived Benefit <--> Perceived Barrier	-.085	0.007	Established
Perceived Benefit <--> Cue to Action	.826	0.682	Not Established
Perceived Benefit <--> Self-Efficacy	.349	0.122	Established
Perceived Barrier <--> Cue to Action	-.001	0.000	Established
Perceived Barrier <--> Self-Efficacy	-.157	0.025	Established
Cue to Action <--> Self-Efficacy	.600	0.360	Established

Internal consistency was deemed to be acceptable if the Cronbach's alpha coefficients were more than 0.7. According to reliability analyses, Cronbach α of the Perceived Susceptibility, Perceived Severity, Perceived Benefits, Perceived Barriers, Cues to Action, Self-Efficacy and Health Behavioural Intention were 0.83, 0.81, 0.86, 0.80, and 0.71, respectively in Phase 1 and 0.87, 0.88, 0.86, 0.89, 0.84, 0.87 and 0.86, respectively in Phase 2.

The detailed mean and standard deviation of each HBM component and Health Behavioural Intention to the socio-demographic characteristics are displayed in Table 5. The total mean score for each domain ranged from 14.03 to 21.32.

Table 5: Result of the minimum and maximum scores, total mean score and SD in Phase 2 (n=171)

Constructs	N	Minimum	Maximum	Mean	SD
Perceived Susceptibility	171	5	25	14.03	4.04
Perceived Severity	171	5	25	16.81	4.44
Perceived Benefits	171	7	20	16.16	2.86
Perceived Barriers	171	6	30	21.32	4.99
Cue to Action	171	9	25	19.97	3.14
Self-Efficacy	171	9	25	18.04	3.87
Health Behavioural Intention	171	5	20	15.64	3.26

Discussion

At the time of when the study was conducted, there were no published papers based on HBM to evaluate cancer health perception among at-risk smokers. Geographics, ethnic groups, national, social,

and genetic-related factors contribute to the disproportionate burden of cancer. Sarawak's population vulnerability to NPC is among the highest in the world. Although genetic predisposition may be the most important risk factor leading to higher incidence of NPC in Sarawak, individual behavioural variables are a key driver of community health that must not be underestimated [35].

A total of 34 items in the questionnaire were formulated consistent with the HBM and divided into two sections: HBM scale for smokers' perception to NPC; and health behavioural intention to smoking cessation. Both sections are constructed with a 5-point Likert scale ranging from "1 = strongly disagree" to "5 = strongly agree". The first section consists of 30 items and is arranged into six subcategories, each representing the six constructs of the health belief model – perceived susceptibility (5 items), perceived seriousness (5 items), perceived benefits (4 items), perceived barriers (6 items), cues to action (5 items), and self-efficacy (5 items). The second section includes 4 items that predict the health behavioural intention (See Table S1 and S2).

Based on the validity and reliability tests, including face and content validity, con-struct validity, and internal consistency, the findings of the current study indicate that the questionnaire has promising psychometric properties. Ten experts were advised on content validity, and 7 items that did not reach the threshold of CVR based on Lawshe's Table and were judged superfluous were removed [25]. Face validity was examined in a pilot study with 10 smokers who fulfilled the eligibility requirements to ensure cultural acceptance and assess relevance and readability within the local community. Given a satisfactory Impact Score of each item, there was no elimination of item in the face validity stage. In the main study, at-risk smokers from various ethnic groups participated, which was conducted in urban, suburban and rural regions of Miri (the northern region of Sarawak). The KMO test yielded a result of 0.697 (Phase 1) and 0.830 (Phase 2) while the Bartlett's test of sphericity obtained 1746.76 (Phase 1) and 3362.86 (Phase 2), both with p-value < 0.001, indicating that the sample size was adequate and the correlation between the items was sufficient for factor analysis.

Construct validity primarily concerns the degree to which a concept measures what it claims to measure [36]. Similar to a previous study [37], a number of analyses were conducted to assess the construct validity, including EFA, CFA, as well as convergent and discriminant validity. The EFA demonstrated in Phase 1 that the seven-factor structure accounted for 63.01 percent of the overall variance. The cut-off point for factor loading was fixed at 0.30. According to EFA suggestion, HBI2 item (I am trying to quit smoking to prevent nose and throat cancer at this time) should be grouped together with Perceived Benefit since the factor loading (0.537) is higher than when it is grouped with Health Behavioural Intention (0.307). This item requires immediate smoking cessation, which may spark inconsistent answers as some respondents may not be prepared to quit. The Cronbach's α coefficient value of 'Health Behavioural Intention' construct appeared to be satisfactory (Phase 1: 0.705; Phase 2: 0.861) and this item could potentially be essential in a larger scale. Thus, the researchers agreed not to delete this item.

In Phase 2, CFA was performed to see if the seven-factor model derived by EFA could demystify the association among the items based on the chosen framework. The loadings of all factors were more than 0.5. The goodness of fit was demonstrated to have an acceptable fit with the data with Chi Square = 641.705; df = 500; P< 0.001; CFI = 0.953; TLI: 0.948; RMSEA= 0.041 (90% CI 0.031 to 0.050). Following CFA, convergent and discriminant validity were tested in Phase 2. The CR and AVE values for each component have to be higher than 0.7 and 0.5; respectively, which are fulfilled in the current study, suggesting an acceptable convergent validity [38]. Discriminant validity was evaluated between the HBM constructs (excluding health behavioural intention construct). To establish an acceptable discriminant validity, the factors' correlation coefficients with other factors must not be greater than each factor's AVE square root [34]. The current findings demonstrated established discriminant

validity for all except for Perceived Benefit <-->Cue to Action. The explanation for this might be that the greater the perceived benefits of quitting smoking, the more likely smokers will look for cues to participate in such health-promoting behaviour, or vice versa. Future studies could delve deeper into the strength of the correlation, particularly between Perceived Benefit and Cue to Action. In terms of reliability, each construct for both Phase 1 and 2 showed rationally acceptable Cronbach's α coefficient values as all of which were higher than 0.7, which demonstrates a high internal consistency [39].

This study has several limitations. Firstly, it is a cross-sectional study using convenience sampling, and thus, it is susceptible to recall and selection bias. Participants, on the other hand, were given ample time to consider their responses before answering the questions. The second limitation was the relatively small sample size for both EFA (n=100) and CFA (n=171). However, Kline (1994) indicated that for EFA, a sample size of 100 is sufficient, while Anderson & Gerbing (1988) suggested that CFA/SEM may be reliably examined with a minimum sample size of 100–150 [40, 41]. The third limitation is the health behavioural intentions investigated in this study were mainly on smoking cessation, the smokers' cultural perspective of other health behavioural intentions was not presented. Future studies looking at different perspective of health behavioural intentions will help to review the psychometric properties of the instrument.

This study provides practical implications. With it being valid and reliable, the public health officials and researchers, now have a reason to launch larger population-based study on health behavioural intention to minimise NPC. Given that the current smoking rates in Malaysia remain high, this questionnaire can help in the understanding and determining the construct that influences smokers' health decisions. Subsequently, cancer risk can be reduced by better prediction, a comprehensive tobacco control programme, policy creation, and health interventions.

Conclusions

The current study developed a comprehensive HBM-based questionnaire with satisfactory psychometric properties, confirming the validity and reliability. A health intervention targeting NPC in this population may be more effective due to the awareness of this population regarding their increased susceptibility to NPC and may also benefit from being informed by this questionnaire. With the possibility of being expanded to general health campaigns that target tobacco smoking, the authors also propose further studies to use the instruments for application in other smoking-related cancers in different susceptible populations and geographic locations.

Funding

Funding for the publication fee was received from RCSI & UCD Malaysia Campus.

Conflicts of Interest

The authors declare that there are no competing or potential conflicts of interest.

Author Contributions

Conceptualization: Kueh Martin, Fairuz Fadzilah Rahim, Abdul Rashid.

Formal analysis: Kueh Martin, Fairuz Fadzilah Rahim.

Investigation: Kueh Martin, Fairuz Fadzilah Rahim.

Methodology: Kueh Martin, Fairuz Fadzilah Rahim.

Project administration: Kueh Martin, Fairuz Fadzilah Rahim.

Resources: Kueh Martin, Fairuz Fadzilah Rahim.

Software: Kueh Martin.

Supervision: Fairuz Fadzilah Rahim, Abdul Rashid.

Validation: Kueh Martin.

Visualization: Kueh Martin.

Writing – original draft: Kueh Martin.

Writing – review & editing: Kueh Martin, Fairuz Fadzilah Rahim, Abdul Rashid.

Institutional Review Board Statement

The study was approved by Joint Penang Independent Ethics Committee of RCSI & UCD Malaysia Campus on 13 August 2020 (Approval no. JPEC 20 0027).

Informed Consent Statement

Informed consent was obtained from all participants involved in the study.

Data Availability Statement

Data supporting the claimed findings can be found at https://datadryad.org/stash/share/QnycI5Un_IWTqoXOr8Zh6xJcsM0RLturj2x7qHAb-ks [30].

Acknowledgments

We express our gratitude to the hospital director in Miri and local estate managers for their helpful assistance. The participants are sincerely appreciated and much acknowledged.

Word count 4292

References

1. Institute for Health Metrics and Evaluation (IHME). Findings from the Global Burden of Disease Study 2017. In: IHME, editor. Seattle, WA. 2018. Available from: http://www.healthdata.org/policy-report/findings-global-burden-disease-study-2017
2. World Health Organization. WHO Report on Cancer: Setting priorities, investing wisely and providing care for all. 2020. Available from: https://apps.who.int/iris/handle/10665/330745
3. Wellmann KF. [Smoking and Health. On the Report of the Advisory Committee to the Surgeon General of the Public Health Service]. Dtsch Med Wochenschr. 1964, 89, 1085-6.
4. Hecht SS. Research opportunities related to establishing standards for tobacco products under the Family Smoking Prevention and Tobacco Control Act. Nicotine & Tobacco Research. 2011, 14(1), 18-28. https://doi.org/10.1093/ntr/ntq216
5. Viswanath K, Herbst RS, Land SR, Leischow SJ, Shields PG. Tobacco and Cancer: An American Association for Cancer Research Policy Statement. Cancer Research. 2010, 70(9), 3419-30. https://doi.org/10.1158/0008-5472.can-10-1087
6. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021, 71(3), 209-49. https://doi.org/10.3322/caac.21660
7. Hashim D, Boffetta P, La Vecchia C, Rota M, Bertuccio P, Malvezzi M, et al. The global decrease in cancer mortality: trends and disparities. Ann Oncol. 2016, 27(5), 926-33. https://doi.org/10.1093/annonc/mdw027
8. Zavala VA, Bracci PM, Carethers JM, Carvajal-Carmona L, Coggins NB, Cruz-Correa MR, et al. Cancer health disparities in racial/ethnic minorities in the United States. Br J Cancer. 2021, 124(2), 315-32. https://doi.org/10.1038/s41416-020-01038-6
9. Lam KO, Lee AW, Choi CW, Sze HC, Zietman AL, Hopkins KI, et al. Global Pattern of Nasopharyngeal Cancer: Correlation of Outcome With Access to Radiation Therapy. Int J Radiat Oncol Biol Phys. 2016, 94(5), 1106-12. https://doi.org/10.1016/j.ijrobp.2015.11.047
10. Azizah A, Hashimah B, Nirmal K, Siti Zubaidah A, Puteri N, Nabihah A, et al. Malaysian National Cancer Registry Report 2012-2016 [Internet]. Malaysia: National Cancer Institute. . Putrajaya, Malaysia: National Cancer Institute, Ministry of Health, 2019. Available from: https://www.moh.gov.my/moh/resources/Penerbitan/Laporan/Umum/2012-2016%20(MNCRR)/MNCR_2012-2016_FINAL_(PUBLISHED_2019).pdf
11. Devi BCR, Pisani P, Tang TS, Parkin DM. High Incidence of Nasopharyngeal Carcinoma in Native People of Sarawak, Borneo Island. Cancer Epidemiology Biomarkers & Prevention. 2004, 13(3), 482-6.
12. Tsao SW, Yip YL, Tsang CM, Pang PS, Lau VM, Zhang G, et al. Etiological factors of nasopharyngeal carcinoma. Oral Oncol. 2014, 50(5), 330-8. https://doi.org/10.1016/j.oraloncology.2014.02.006
13. Linton RE, Daker M, Khoo ASB, Choo DCY, Viljoen M, Neilsen PM. Nasopharyngeal carcinoma among the Bidayuh of Sarawak, Malaysia: history and risk factors (Review). Oncol Lett. 2021, 22(1), 514. https://dx.doi.org/10.3892%2Fol.2021.12775
14. Guo S-S, Huang P-Y, Chen Q-Y, Liu H, Tang L-Q, Zhang L, et al. The impact of smoking on the clinical outcome of locoregionally advanced nasopharyngeal carcinoma after chemoradiotherapy. Radiation Oncology. 2014, 9(1), 246. https://doi.org/10.1186/s13014-014-0246-y
15. Xue WQ, Qin HD, Ruan HL, Shugart YY, Jia WH. Quantitative association of tobacco smoking with the risk of nasopharyngeal carcinoma: a comprehensive meta-analysis of studies conducted between 1979 and 2011. Am J Epidemiol. 2013, 178(3), 325-38. https://dx.doi.org/10.1093%2Faje%2Fkws479
16. Institute for Public Health. National Health and Morbidity Survey 2015 (NHMS 2015). Ministry of Health Malaysia Kuala Lumpur; 2015. Available from: https://www.moh.gov.my/moh/resources/nhmsreport2015vol2.pdf

17. Rahman M, Arif M, Razak MFA, Suhaili M, Tambi Z, Akoi C, et al. Factors associated with tobacco use among the adult population in Sarawak, Malaysia: A cross sectional study. Epidemiology Biostatistics and Public Health. 2015, 12, 1-9. https://doi.org/10.2427/10292
18. Champion VL, Skinner CS. The health belief model. Health behavior and health education: Theory, research, and practice, 4th ed. San Francisco, CA, US: Jossey-Bass, 2008. p. 45-65.
19. Maiman LA, Becker MH. The Health Belief Model: Origins and Correlates in Psychological Theory. Health Education Monographs. 1974, 2(4), 336-53. https://doi.org/10.1177/109019817400200404
20. Odedina FT, Dagne G, Pressey S, Odedina O, Emanuel F, Scrivens J, et al. Prostate cancer health and cultural beliefs of black men: The Florida Prostate Cancer Disparity Project. Infectious Agents and Cancer. 2011, 6(2), S10. https://dx.doi.org/10.1186%2F1750-9378-6-S2-S10
21. Solomon K, Tamire M, Kaba M. Predictors of cervical cancer screening practice among HIV positive women attending adult anti-retroviral treatment clinics in Bishoftu town, Ethiopia: the application of a health belief model. BMC Cancer. 2019, 19(1), 989. https://doi.org/10.1186/s12885-019-6171-6
22. Tavafian SS, Hasani L, Aghamolaei T, Zare S, Gregory D. Prediction of breast self-examination in a sample of Iranian women: an application of the Health Belief Model. BMC Women's Health. 2009, 9(1), 37. https://doi.org/10.1186/1472-6874-9-37
23. LaMorte., Wayne. W. The Health Belief Model: Behavioral Change Models 2019. Available from: http://sphweb.bumc.bu.edu/otlt/MPH-Modules/SB/BehavioralChangeTheories/BehavioralChangeTheories2.html
24. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine. 2000, 25(24), 3186-91. https://doi.org/10.1097/00007632-200012150-00014
25. Lawshe CH. A quantitative approach to content validity. Personnel psychology. 1975, 28(4), 563-75. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x
26. Samet JM, Wipfli H, Platz EA, Bhavsar N. A Dictionary of Epidemiology, Fifth Edition: Edited by Miquel Porta. American Journal of Epidemiology. 2009, 170(11), 1449-51. https://doi.org/10.1093/aje/kwp322
27. Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. Frontiers in public health, 6, 149. https://doi.org/10.3389/fpubh.2018.00149
28. Lacasse Y, Godbout C, Sériès F. Health-related quality of life in obstructive sleep apnoea. Eur Respir J. 2002, 19(3), 499-503. https://doi.org/10.1183/09031936.02.00216902
29. Suhr DD. Exploratory or confirmatory factor analysis? 2006. Available from: https://support.sas.com/resources/papers/proceedings/proceedings/sugi31/200-31.pdf
30. [dataset] Kueh M, Rahim FF, Rashid A. Data from: Development and validation of the health behavioural intention on smoking cessation to prevent nasopharyngeal cancer in Sarawak, Malaysia based on the Health Belief Model: a cross-sectional study. Dryad Digital Repository, November 18, 2021. https://doi.org/10.5061/dryad.4tmpg4fbb
31. Munro BH. Statistical methods for health care research: lippincott williams & wilkins; 2005.
32. Xiang B, Wong HM, Cao W, Perfecto AP, McGrath CPJ. Development and validation of the Oral health behavior questionnaire for adolescents based on the health belief model (OHBQAHBM). BMC Public Health. 2020, 20(1), 701. https://doi.org/10.1186/s12889-020-08851-x
33. Brown TA. Confirmatory factor analysis for applied research. New York, NY, US: The Guilford Press; 2006.
34. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. Journal of marketing research. 1981, 18(1), 39-50. https://doi.org/10.2307/3151312
35. McGinnis JM, Williams-Russo P, Knickman JR. The case for more active policy attention to health promotion. Health affairs. 2002, 21(2), 78-93. https://doi.org/10.1377/hlthaff.21.2.78

36. Cohen RJ, Swerdlik ME, Phillips SM. Psychological testing and assessment: An introduction to tests and measurement: Mayfield Publishing Co; 1996.
37. Shamsalinia A, Moradi M, Farahani MA, Masoudi R, Ghadimi R, Rad RE, et al. Designing and psychometric evaluation of disease-related fear scale (D-RFS) in adults with epilepsy: A sequential exploratory mixed methods design. Epilepsy & Behavior. 2020, 110, 107169. https://doi.org/10.1016/j.yebeh.2020.107169
38. Schreiber JB, Nora A, Stage FK, Barlow EA, King J. Reporting structural equation modeling and confirmatory factor analysis results: A review. The Journal of educational research. 2006, 99(6), 323-38. https://doi.org/10.3200/JOER.99.6.323-338
39. Kline P. The handbook of psychological testing: Psychology Press, 2000. https://doi.org/10.4324/9781315812274
40. Kline P. An Easy Guide to Factor Analysis (1st ed.). 1st edition ed. London. 1994, p. 208. https://doi.org/10.4324/9781315788135
41. Anderson JC, Gerbing DW. Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. Psychological Bulletin. 1988, 103, 411-423. http://dx.doi.org/10.1037/0033-2909.103.3.411

Table S1: Questionnaire to predict health behavioural intention on smoking cessation to prevent nasopharyngeal cancer in Sarawak, Malaysia based on the Health Belief Model: a cross-sectional study (English Version)

Section 1: Perceptions towards nose and throat cancer

Please choose the best answer options to indicate your level of agreement.

Constructs of Health Belief Model	Questions	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Perceived susceptibility	1. I feel that I will get nose and throat cancer in the future.					
	2. My chances of getting nose and throat cancer are high.					
	3. I cannot avoid myself from getting nose and throat cancer.					
	4. I am worried about getting nose and throat cancer due to family history.					
	5. My smoking habit makes me more likely than average to get nose and throat cancer.					
Perceived seriousness	6. The thought of nose and throat cancer scares me.					
	7. If I had nose and throat cancer, the cost of treatment can be a financial strain.					
	8. Nose and throat cancer would threaten a relationship with my family.					

	9. If I had nose and throat cancer my whole life would change.					
	10. If I developed nose and throat cancer, I would not live long.					
Perceived benefits	11. Quit smoking decreases my chance of getting nose and throat cancer.					
	12. Quit smoking decreases my chance of dying from nose and throat cancer.					
	13. Quit smoking can improve my health.					
	14. I feel less anxious about nose and throat cancer if I quit smoking.					
Perceived barriers	15. It is difficult to quit tobacco smoking (e.g. peer pressure).					
	16. I feel anxious without smoking.					
	17. Tobacco smoking relieves my stress.					
	18. I experience headache without smoking.					
	19. I experience excessive salivation without smoking.					
	20. I think getting nose and throat cancer is destiny and quitting smoking will not change it.					
Cues to action	21. I will stop smoking if I have social support.					

	22. I will stop smoking if there are information sources that reminds me. Examples of sources include: the internet, newspapers, radio and TV.					
	23. I will stop smoking if I have the will to change.					
	24. I will stop smoking if I know the diseases related to smoking.					
	25. I will stop smoking if there are health professionals to assist me.					
Self-efficacy	26. I can refuse to smoke when I am thinking about difficult problem.					
	27. I can refuse the urge to smoke.					
	28. I can refuse to smoke when I see someone else smoking.					
	29. I can refuse to smoke when offered by my friends/ family.					
	30. I can refuse from buying cigarettes when I have extra pocket money.					

Section 2: Health Behavioural Intention to quit smoking

Please choose the best answer options to indicate your level of agreement.

Statements	Strongly disagree	Disagree	Neutral	Agree	Strongly agree

1. I would like to lead a healthier lifestyle to prevent nose and throat cancer.					
2. I am trying to quit smoking to prevent nose and throat cancer at this time.					
3. I plan to quit smoking to prevent nose and throat cancer within six months.					
4. I would like to quit smoking to prevent nose and throat cancer but have never really tried.					

Table S2: Questionnaire to predict health behavioural intention on smoking cessation to prevent nasopharyngeal cancer in Sarawak, Malaysia based on the Health Belief Model: a cross-sectional study (Malay Version)

Bahagian 1: Persepsi terhadap kanser hidung dan tekak.

Sila bulatkan dengan pilihan jawapan terbaik untuk menunjukkan tahap persetujuan anda.

Domain Model Kepercayaan Kesihatan	Soalan	Sangat tidak setuju	Tidak setuju	Kurang bersetuju	Setuju	Sangat setuju
Kepercayaan kepada tiada daya tahan	1. Saya merasakan bahawa saya akan mendapat kanser pangkal hidung pada masa akan datang.					
	2. Peluang saya mendapat kanser pangkal hidung adalah tinggi.					
	3. Saya tidak dapat mengelakkan diri daripada mendapat kanser pangkal hidung.					
	4. Saya bimbang terkena kanser pangkal hidung kerana ahli keluarga saya menghadapi masalah ini.					
	5. Tabiat merokok saya menyebabkan saya lebih berisiko menghidap kanser pangkal hidung.					
Kepercayaan kepada bahaya	6. Risiko dan kesan-kesan kanser pangkal hidung menakutkan saya.					
	7. Jika saya mempunyai kanser pangkal hidung, kos					

	rawatan akan menjadi beban kewangan keluarga saya.					
	8. Kanser pangkal hidung akan merosakkan hubungan antara ahli keluarga saya.					
	9. Jika saya mempunyai kanser pangkal hidung, kehidupan saya akan diubah.					
	10. Jika saya mempunyai kanser pangkal hidung, saya tidak akan hidup lama.					
Kepercayaan kepada manfaat	11. Berhenti merokok akan mengurangkan peluang saya mendapat kanser pangkal hidung.					
	12. Berhenti merokok akan mengurangkan risiko kematian disebabkan kanser pangkal hidung.					
	13. Berhenti merokok amat penting untuk meningkatkan kesihatan saya.					
	14. Berhenti merokok akan mengurangkan kerisauan saya terhadap kanser pangkal hidung.					
Kepercayaan kepada halangan	15. Saya berasa sukar untuk menghentikan tabiat merokok atas					

	sebab tekanan rakan sebaya dan lain-lain.					
	16. Saya akan berasa bimbang akan sesuatu tanpa merokok.					
	17. Merokok menghilangkan tekanan saya.					
	18. Jika saya tidak merokok, saya akan mengalami sakit kepala					
	19. Jika saya tidak merokok, saya akan mengalami air liur berlebihan.					
	20. Saya rasa mendapat kanser pangkal hidung adalah takdir dan berhenti merokok tidak akan mengubahnya.					
Isyarat untuk bertindak	21. Jika saya mempunyai sokongan sosial, saya akan berhenti merokok.					
	22. Jika beberapa sumber maklumat mengingatkan saya, saya akan berhenti merokok. Contoh sumber maklumat termasuk internet, surat khabar, radio dan TV.					
	23. Saya akan berhenti merokok sekiranya saya mempunyai kemahuan untuk berubah.					

	24. Menyedari bahaya merokok akan membantu saya berhenti merokok.					
	25. Saya akan berhenti merokok sekiranya ahli profesional kesihatan membantu saya					
Keberkesanan diri	26. Saya boleh menolak merokok apabila saya ingin memikirkan masalah yang sukar.					
	27. Saya boleh menolak merokok apabila saya mempunyai keinginan untuk merokok.					
	28. Saya boleh menolak merokok apabila saya melihat orang lain merokok.					
	29. Saya boleh menolak merokok apabila kawan/ keluarga saya mengajak saya untuk merokok.					
	30. Saya boleh menolak membeli rokok apabila saya mempunyai lebihan wang saku.					

Bahagian 2: Tingkah laku untuk berhenti merokok demi kesihatan

Sila bulatkan dengan pilihan jawapan terbaik untuk menunjukkan tahap persetujuan anda.

Soalan	Sangat tidak bersetuju	Tidak setuju	Kurang bersetuju	Setuju	Sangat setuju

1. Saya akan menjalani gaya hidup yang lebih sihat untuk mencegah kanser pangkal hidung.					
2. Saya sedang berusaha untuk berhenti merokok untuk mengurangkan risiko kanser pangkal hidung pada masa ini.					
3. Saya merancang untuk berhenti merokok untuk mengurangkan risiko kanser pangkal hidung dalam masa enam bulan.					
4. Saya ingin berhenti merokok untuk mengurangkan risiko kanser pangkal hidung tetapi tidak pernah mencuba.					

Reporting checklist for cross sectional study.

Based on the STROBE cross sectional guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the STROBE cross sectionalreporting guidelines, and cite them as:

von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies.

		Reporting Item	Page Number
Title and abstract			
Title	#1a	Indicate the study's design with a commonly used term in the title or the abstract	3
Abstract	#1b	Provide in the abstract an informative and	3

		balanced summary of what was done and what was found	
Introduction			
Background / rationale	#2	Explain the scientific background and rationale for the investigation being reported	4-5
Objectives	#3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	#4	Present key elements of study design early in the paper	6
Setting	#5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6
Eligibility criteria	#6a	Give the eligibility criteria, and the sources and methods of selection of participants.	6
	#7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6-7
Data sources / measurement	#8	For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Give information separately for for	7-8

		exposed and unexposed groups if applicable.	
Bias	#9	Describe any efforts to address potential sources of bias	7-8. We involved experts and participants of different background and ethnic groups in hopes to reduce biasness.
Study size	#10	Explain how the study size was arrived at	14
Quantitative variables	#11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	8
Statistical methods	#12a	Describe all statistical methods, including those used to control for confounding	7-8
Statistical methods	#12b	Describe any methods used to examine subgroups and interactions	8
Statistical methods	#12c	Explain how missing data were addressed	N/A. No missing data was collected.
Statistical methods	#12d	If applicable, describe analytical methods taking account of sampling strategy	8
Statistical methods	#12e	Describe any sensitivity analyses	N/A
Results			
Participants	#13a	Report numbers of individuals at each stage of	N/A. Do not have a list of

		study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed. Give information separately for for exposed and unexposed groups if applicable.	numbers potentially eligible.
Participants	#13b	Give reasons for non-participation at each stage	N/A
Participants	#13c	Consider use of a flow diagram	N/A. Do not have a list of numbers potentially eligible.
Descriptive data	#14a	Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Give information separately for exposed and unexposed groups if applicable.	8-9
Descriptive data	#14b	Indicate number of participants with missing data for each variable of interest	N/A. No missing data was collected
Outcome data	#15	Report numbers of outcome events or summary measures. Give information separately for exposed and unexposed groups if applicable.	N/A. This study focuses on validating the questionnaire with a series of analysis.
Main results	#16a	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their	N/A. This study focuses on validating the

		precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	questionnaire with a series of analysis.
Main results	#16b	Report category boundaries when continuous variables were categorized	N/A. This study focuses on validating the questionnaire with a series of analysis.
Main results	#16c	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A. This study focuses on validating the questionnaire with a series of analysis.
Other analyses	#17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	10-13
Discussion			
Key results	#18	Summarise key results with reference to study objectives	13-14
Limitations	#19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.	14
Interpretation	#20	Give a cautious overall interpretation considering objectives, limitations, multiplicity	14

		of analyses, results from similar studies, and other relevant evidence.	
Generalisability	#21	Discuss the generalisability (external validity) of the study results	14
Other Information			
Funding	#22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	14

Notes:

- 9: 5 - 6. We involved experts and participants of different background and ethnic groups in hopes to reduce biasness.
- 12c: N/A. No missing data was collected.
- 13a: N/A. Do not have a list of numbers potentially eligible.
- 13c: N/A. Do not have a list of numbers potentially eligible.
- 14b: N/A. No missing data was collected
- 15: N/A. This study focuses on validating the questionnaire with a series of analysis.
- 16a: N/A. This study focuses on validating the questionnaire with a series of analysis.
- 16b: N/A. This study focuses on validating the questionnaire with a series of analysis.

- 16c: N/A. This study focuses on validating the questionnaire with a series of analysis. The STROBE checklist is distributed under the terms of the Creative Commons Attribution License CC-BY. This checklist was completed on 19. September 2021 using https://www.goodreports.org/, a tool made by the EQUATOR Network in collaboration with Penelope.ai

BMJ Open

Development and validation of the Health Belief Model questionnaire to promote smoking cessation for nasopharyngeal cancer prevention: a cross-sectional study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-057552.R2
Article Type:	Original research
Date Submitted by the Author:	07-Jul-2022
Complete List of Authors:	Kueh, Martin; RCSI & UCD Malaysia Campus, Department of Public Health Medicine Rahim, Fairuz Fadzilah; RCSI & UCD Malaysia Campus, Department of Public Health Medicine Rashid, Abdul; RCSI & UCD Malaysia Campus, Public Health Medicine
Primary Subject Heading:	Public health
Secondary Subject Heading:	Public health, Smoking and tobacco, Oncology
Keywords:	PUBLIC HEALTH, Head & neck tumours < ONCOLOGY, Epidemiology < ONCOLOGY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Development and validation of the Health Belief Model questionnaire to promote smoking cessation for nasopharyngeal cancer prevention: a cross-sectional study

Martin Kueh[1], Fairuz Fadzilah Rahim[1] & Abdul Rashid[1]

[1] Department of Public Health Medicine, RCSI & UCD Malaysia Campus, No.4, Jalan Sepoy Lines, 10450 George Town, Penang Malaysia

Correspondence: Martin Kueh, Department of Public Health Medicine, RCSI & UCD Malaysia Campus, No.4, Jalan Sepoy Lines, 10450 George Town, Penang Malaysia. Tel: 6016-879 2269. Email: martin1kueh@hotmail.com

Abstract

Objective Lifestyle-induced nasopharyngeal carcinoma (NPC) is a serious but preventable risk factor. This study serves to develop and validate a questionnaire that aims to predict the health behavioural intention on smoking cessation in Sarawak, Malaysia using the Health Belief Model (HBM).

Design A cross-sectional study

Setting Urban and suburban areas in Sarawak, Malaysia.

Participants The preliminary items of the instrument were developed after extensive literature review. The instrument was translated into the Malay language using the forward-backwards method before commencing with the content validity by a panel of 10 experts. Face validity was done both quantitatively and qualitatively by 10 smokers. The construct validity of the instrument was evaluated through exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). A total of 100 smokers participated in phase 1 for EFA, while 171 smokers participated in phase 2 for CFA. Internal consistency was measured using Cronbach's alpha coefficients to evaluate the reliability.

Results In the exploratory stage, the factor loading of each item remained within the acceptable threshold. The final revised CFA yielded appropriate fit of the seven-factor model with the following model fit indexes: Chi-Square: 641.705; df= 500; P< 0.001; CFI = 0.953; TLI: 0.948; RMSEA= 0.041. Satisfactory convergent validity and divergent validity were shown, with the exception of one pairwise construct. The internal reliability of these scales was above the desirable threshold, with Cronbach's alpha coefficients ranging from 0.705 to 0.864 and 0.838 to 0.889 in phase 1 and 2, respectively.

Conclusions The study substantiated the instrument to be valid and reliable for predicting smokers' health behavioural intention to reduce cancer risk. The instrument is made up of 34 items, categorised into two sections, six HBM constructs and health behavioural intention. The instrument can be utilised for other smoking cessation-related cancers in different at-risk populations.

Keywords: tobacco smoking, cancer, health promotion, nasopharyngeal cancer, health belief model, development, validation, smoking cessation

Article Summary

Strengths and limitations of this study

1. This study established a novel instrument to assess smokers' health behavioural intention to reduce NPC risks, based on a well-known framework in a series of systematic validation stages. This instrument can potentially be applied to other smoking-related cancers as well.

2. Face validity was undertaken both qualitatively and quantitatively to sufficiently reflect the demographic during the assessments of psychological constructs. The validation was markedly aided by experts' evaluations.

3. This study was conducted in two phases, involving both urban and suburban smokers, to examine concept validity, convergent validity, and divergent validity.

4. The study's generalizability may be limited by the smokers' cultural perspective, hence, further studies on smokers from different cultures will be needed to assess the instrument's psychometric properties.

Introduction

According to the World Health Organization, tobacco smoking is a public health concern that accounts for over 8 million deaths per year and is the leading avoidable cause of illness, disability, and death globally [1]. Annually, exposure to smoking is associated with 2.4 million deaths from cancer throughout the world [2]. The report from the Surgeon General of the United States of America (USA) associating smoking and cancer was a watershed moment in public health towards tobacco's adverse effects on human health. This was followed by the subsequent discovery that tobacco smoke comprises approximately 7,000 compounds, 72 of which are carcinogenic [3, 4]. Tobacco use is now causally associated with at least 20 cancer types. There are wide-ranging immediate and long-term health benefits accompanying smoking cessation [2]. However, the harmful consequences of tobacco smoking are widely neglected or underestimated, despite the fact that it remains a significant public health hazard among the impoverished, and marginalised, as well as those in developing nations, which bear a disproportionate share of the burden [5].

Cancer is a leading cause of death as well as a major obstacle to improving lifespan in every country [6]. There is an approximately 1% reduction in the overall cancer mortality rate across both sexes in both high- and low-income countries [7]. However, there are variations in the frequency and distribution of aetiological aspects such as socioeconomic, geographical, genetic, biological, ethnic, social, and physical factors, as well as disparities across cancer types [7, 8]. For instance, nasopharyngeal cancer (NPC) is uncommon but unique among head and neck cancers with its own distinct epidemiological & risk factors. Global data from the World Health Organization illustrates poorer outcomes of NPC in endemic areas like Southeast Asia which has an unbalanced global burden of 67% [9].

In Malaysia, NPC is a nationwide public health concern and the fifth leading form of cancer, amounting to 4,597 new diagnoses of NPC for the 2012-2016 period. A recent report from the Malaysian National Cancer Registry reported that the lifetime risk of developing NPC among men and women is 1 in 175 and 1 in 482, respectively [10]. There is a substantial geographical variance within the country, with Sarawak exhibiting a higher prevalence rate of NPC. A previous study has shown a significant high age-standardised rate in males (13.5/100,000, 95% Confidence Interval = 12.2 – 15.0) and females

(6.2/100,000, 95% CI= 5.7-6.7). The high-risk ethnic groups in Sarawak which include Bidayuh, Chinese, Iban, Malays and Melanau collectively rank top globally. In particular, the risk among the Bidayuh ethnic population, which is a native indigenous group, exceeds the risk for the male and female general population in Sarawak by 2.3 times and 1.9 times respectively [11].

This trend has been ascribed to potential risk factors, which include Epstein-Barr virus, genetic susceptibility, consumption of food with nitrous compounds and volatile nitrosamines, and complex interaction with environmental factors [12]. Among the many risk factors that are associated with NPC, tobacco smoking is the most important modifiable cause of NPC [13, 14]. A meta-analysis 32 epidemiological studies (28 case-control studies and 4 cohort studies) on the association of tobacco smoking and NPC from 1979 to 2011 reported that tobacco-correlated NPC cases were 60% higher compared to non-smokers [15]. The Malaysian National Health and Morbidity Survey 2015 showed that the prevalence of tobacco smoking among the population in Sarawak was 25.4%. The native indigenous male (61.2%) and female (10.7%) smokers in Sarawak were among the highest nationwide [16]. Rahman et al (2015) indicated that the average number of cigarettes smoked per day in Sarawak was 13.6 [17]. This lifestyle-induced NPC is a serious concern, and necessitates a preventative strategy centred on modifying health risk behaviours.

Despite the robust establishment of cumulative impact of tobacco smoking on the risk of cancer, in Malaysia, there is still a paucity of published studies and research to evaluate the effectiveness of comprehensive strategies to promote cancer prevention among smokers. It is imperative to create a questionnaire that focuses on behavioural factors and is customised to the interests of local smokers to risk of cancer. The Health Belief Model (HBM) is an approachable theoretical model that aids in the understanding of the individual's or a smoker's belief on the health-related behavioural intention [18, 19]. This HBM can predict the smokers' effort to improve health or their health-seeking behaviours for the prevention of NPC. This study serves to develop and validate a health behaviour model (HBM)-based questionnaire that aims to predict the health behavioural intention of smokers and their perspective and motivation towards smoking cessation.

Conceptual Framework

The Health Belief Model (HBM) is the underpinning conceptual framework for this study to focus on psychological variables to predict health behavioural intention. Developed in the 1950s by social psychologists at the United States Public Health Service, the Health Belief Model (HBM) is currently one of the most extensively used cognitive model and theoretical framework to help researchers understand and predict health behaviours in the population and ultimately guide health promotion and interventions activities [18, 19]. A large volume of studies conducted in numerous countries, both in developed and developing countries, have utilized HBM to examine health-promoting behaviours for the prevention of different cancers [20-22]. The HBM is a value-expectancy theory, based on the hypothesis of Lewin et al that highlights the influence of two variables on behaviour: 1) the value that a person places on the outcome of the behaviour and 2) the person's perception of how likely the behaviour will lead to that outcome, in the event of an illness (19). Having evolved over the past decades, HBM currently consists of 6 elements: i) perceived susceptibility, ii) perceived severity, iii) perceived benefits, iv) perceived barriers, v) cues to action, and vi) self-efficacy [23].

The first four refer to a person's subjective perceptions regarding 1) his/her risk of getting the disease; 2) how severe the consequences are of getting the disease; 3) the benefits from performing a health behaviour in preventing, curing, or managing the disease; and 4) obstacles to that health behaviour, e.g. financial and time costs, side effects, and so on (23). "Cue to action" is the stimulus, which may be internal (e.g., physical sensations) or external (e.g., friends with the disease, social media), that is required for that health behaviour to occur, and "self-efficacy" refers to the person's confidence on how capable he/she is to successfully undertake that health behaviour [19, 23].

Methods

Study design and setting

A cross-sectional study was conducted in Sarawak, Malaysia in two phases: phase 1 from October 2020 to January 2021 and phase 2 from January to April 2021. Sarawak is the largest Malaysian state situated on the island of Borneo with a population of more than 2.6 million, made up of 26 different ethnic groups. Sarawak is divided into twelve divisions, each of which is further divided into districts and sub-districts. The two divisions selected for the study were Miri and Bintulu, which had the respective populations of 433,800 and 266,200.

The sample population in phase 1 were residents residing in urban and suburban areas in Miri, Sarawak. Phase 2 mainly involved residents of Bintulu, Kuala Baram (a federal constituency in Miri Division), and remote rural areas in Miri. They were mainly local employees working in the agricultural industry. Data were collected by eight trained research assistants. All research assistants were given a crash course in the research aims, methodology and data collection, as well as a trial run to simulate real-world situations. Prior to the distribution of the questionnaires, participants were briefed regarding the objective and methodology of the study, as well as the benefits and risks involved. Involvement in this study was entirely voluntary. Upon clarifying the study details, informed consent was obtained from each participant. The anonymity of the respondents' details is assured. The research assistants provided clarification to smokers who requested assistance. The methodology of the study and data collection were recorded precisely and accurately throughout the process of the research.

The inclusion criteria were: 1) 18 years old and above, 2) smoked for at least a year, and 3) is a Sarawakian. Participants who did not consent to participate, were pregnant, or smoked e-cigarettes only were excluded from this study.

Patient and Public Involvement

No patient was involved in this study.

Instrument development and validation

The questionnaire was self-administered based on prior validated studies. A search was conducted using the National Library of Medicine (PubMed), Google Scholar and Cochrane Library databases by exploring various keywords: health belief model, nasopharyngeal cancer, cancer, smokers, smoking behaviours/habit, and questionnaire/tools/instruments. Questionnaires and prior literature were examined primarily, and explicitly on smokers around NPC prevention using HBM. The initial questionnaire consisted of 41 items.

The questionnaire was developed in English and translated to Malay by two local fluent bilingual translators. An experienced researcher, whose mother tongue is Malay, compared the Malay version of the questionnaire to the English version. A 'back-translation' approach to English was taken independently by another two bilingual translators based on the Beaton-recommended guidelines [24].

To determine the content validity, 10 healthcare professionals including public health experts, hospital directors, and health officers were invited to evaluate the survey instrument. Using Lawshe's model, a questionnaire was designed and organized to assist and allow the panellists to express clearly their views on the importance of including different components in a model. Experts received attachments of the questionnaire via email, which was graded on a three-point scale: essential, useful but not essential, and not essential. Based on Lawshe's table, items with a CVR value greater than 0.62 were retained [25]. To minimise ambiguity, the experts evaluated each item's accuracy, phrasing and grammar as well as their relevance to the construct. Modifications were made for subsequent analysis based on the experts' comments and suggestions. Overall, 8 items were deleted, 2 items were rephrased and 1 item was allocated into a different construct.

Face validity was evaluated through a pilot study by 10 local smokers of different ethnic backgrounds, both qualitatively and quantitatively. Smokers who participated in the pilot study were exempted from the main study. The pilot study is conducted on a small-scale basis to ascertain the feasibility of the proposed larger study [26]. In the qualitative stage, cognitive interviews were conducted face-to-face individually to obtain participants' feedback on their comprehension and answers. Items which were not well understood were identified from the cognitive interview [27]. Minor revisions were made to better suit a linguistically and culturally diverse context. Based on the smokers' perspective, an item-oriented person's faith or spirituality to own health was also included. With the consensus of the researchers, an item was added, 'I think getting nose and throat cancer is my destiny and quitting smoking will not change it'. In the quantitative stage, a survey was disseminated based on a Likert scale of 1 (least importance) to 5 (extremely important) to determine the clinical impact of each item. The importance score was calculated based on the "clinical impact method" in which the clinical impact of each item was determined from the proportion of participants who identified it as important. This technique was chosen for better clarity where the items were ranked according to their impact score. The mean importance score of each item was computed using the following formula: Impact Score = Frequency (Proportion) x Importance. Factors were kept if the Impact Score equal or more than 1.5. These factors were defined as deemed suitable and kept for further evaluation [28]. In the current study, the impact score for each item ranged from 1.7 to 4.6, therefore, no item was eliminated.

Subsequently, in the main study, to determine the construct validity, exploratory and confirmatory factor analyses were performed in two periods. In the first period, EFA (n = 100) was used to determine the number of latent factors or the relationships between the common factors. The model was later adjusted in the second period with CFA (n = 171) via structural equations using AMOS. CFA confirmed the overall fit of the model and indicated that the measures were in the acceptable range [29]. Convergent validity and discriminant validity were also carried out in phase 2. The flow diagram for the questionnaire development and validation is shown in Figure 1.

Final scoring of the instrument:

In the main study, the total mean score for each HBM component was formulated on the five-point Likert scale options from strongly disagree to strongly agree, with scores ranging from 1 to 5 points. With the exception of 'Perceived Barriers,', which is inversely proportional, a greater score reflects a firmer desire to quit smoking.

Ethical considerations:

This study was approved by the Joint Penang Ethics Committee (Approval no. JPEC 20 0027). The informed consent of all participants was obtained voluntarily, and data confidentiality and storage are assured.

Data management:

The data was processed in Microsoft Excel before being analysed in SPSS 26.0 and AMOS 23.0 [30]. Listwise deletion approach was done for missing data of less than 2%. Socio-demographic characteristics are presented as number and percentage distribution. Cronbach's alpha coefficients were used to determine internal consistency. Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) were used to test the construct validity of each construct. Specifically, EFA was evaluated in phase 1 and CFA in phase 2. A p-value below 0.05 was deemed to be statistically significant.

EFA was performed in phase 1 to reveal the fundamental structure of a large set of variables [29]. The factors were extracted using principal component analysis with a varimax rotation. The Kaiser-Meyer-Olkin (KMO) > 0.6 and Bartlett's test for sphericity ($p < 0.05$) were used for adequacy and item checking. All loading factors below 0.3 were excluded from the constructs [31].

CFA was performed in phase 2 to assess the data integrity and the structural model [29]. The acceptable level of standardised factor loading was set at 0.5 and above to ensure a satisfactory association between items and corresponding factors [32]. Different fit indices were utilised to estimate the model fit. These include a comparative fit index (CFI) of > 0.90, Tucker-Lewis Index (TLI) > 0.90 and root mean square error of approximation (RMSEA) < 0.05 [33]. Additionally, convergent validity and discriminant validity were evaluated based on the composite reliability (CR) and average variance extracted (AVE). The HBM components were analysed with the minimum and maximum scores, total mean score and standard deviation (SD).

Results:

A total of 100 and 171 smokers participated in phase 1 and 2 of the study, respectively. The response rate for phase 1 was 100%, whereas in phase 2 the response rate was 98.3% (171/174). The majority of the participants were males (Phase 1, 86.0%; Phase 2, 89.5%) and were in the 30-39 age group (Phase 1, 40.0%; Phase 2, 40.9%). A little over a third of the participants smoked for more than 10 years for both phase 1 (37%) and phase 2 (40.4%). In phase 1, most of the participants were Iban (34%), followed by Malay (19%), Chinese (18%), Others (13%), Melanau (9%) and Bidayuh (7%). Most of the participants in phase 2 were Chinese (33.9%), followed by Iban (24.6%), Malay (12.3%), Bidayuh (9.9%), Melanau (9.9%) and Others (9.4%). Details of the smokers' demographic information is presented in Table 1.

Table 1: Demographic characteristics of participants

	Phase 1 (n=100)		Phase 2 (n=171)	
	Number	Percentage (%)	Number	Percentage (%)
Age				
18-29	34	34.0	46	26.9
30-39	40	40.0	70	40.9
40-49	14	14.0	43	25.1
50-64	12	12.0	10	5.8
65 and above	0	0.0	2	1.2
Gender				
Male	86	86.0	153	89.5
Female	14	14.0	18	10.5

Ethnic groups				
Malay	19	19.0	21	12.3
Chinese	18	18.0	58	33.9
Bidayuh	7	7.0	17	9.9
Iban	34	34.0	42	24.6
Melanau	9	9.0	17	9.9
Others	13	13.0	16	9.4
Years of smoking				
1-5 years	32	32.0	44	25.7
6-10 years	31	31.0	58	33.9
More than 10 years	37	37.0	69	40.4

In phase 1, Kaiser-Meyer-Olkin Measure of Sampling Adequacy was 0.697 and Barlett's test of the sphericity was significant (x2 = 1,746, p-value < 0.001). EFA was conducted to analyse the factor structure with principal component analysis with a varimax rotation. A decision was made to go for 7-factor structures since there is clarity of 7 constructs. The EFA found that 7 variables had eigenvalues larger than Kaiser's threshold of 1 and explained 63.0% of the variance when combined. Factor loadings higher than 0.3 were found in all the items. Four items had cross-loading with values greater than 0.3, which are PBar5, HBI2, HBI3 & HBI4. All items remained because the contents of the items were regarded as relevant based on the decision and judgment of the researchers. Table 2 shows the EFA with total items and the factor loading of each construct for the 7-factor model.

Table 2: Result of Exploratory Factor Analysis in phase 1 (n = 100)

		Component						
Constructs	Items	1	2	3	4	5	6	7
Perceived Susceptibility	PSus1			0.665				
	PSus2			0.750				
	PSus3			0.810				
	PSus4			0.548				
	PSus5			0.736				
Perceived Severity	PSev1						0.558	
	PSev2						0.720	
	PSev3						0.744	
	PSev4						0.843	
	PSev5						0.730	
Perceived Benefit	PBen1	0.767						
	PBen2	0.765						
	PBen3	0.763						
	PBen4	0.792						
Perceived Barrier	PBar1					0.594		
	PBar2					0.725		
	PBar3					0.658		
	PBar4					0.709		
	PBar5					0.590		
	PBar6					0.617		
Cue to action	CUE1				0.721			

	CUE2				0.584			
	CUE3				0.736			
	CUE4				0.723			
	CUE5				0.745			
Self-efficacy	EFF1		0.721					
	EFF2		0.730					
	EFF3		0.770					
	EFF4		0.868					
	EFF5		0.625					
Health Behavioural Intention	HBI1							0.681
	HBI2							0.307
	HBI3							0.733
	HBI4							0.596
Rotation Sums of Squared Loading								
Total		3.469	3.371	3.246	3.072	3.071	3.012	2.185
Percentage of Variance		10.203	9.913	9.548	9.034	9.032	8.859	6.425
Cumulative percentage		10.203	20.116	29.664	38.698	47.730	56.588	63.014

CFA was performed in phase 2 to assess whether the seven-factor model indicated by the EFA could sufficiently represent the data. The items in their respective constructs were loaded between 0.586 and 0.898 (Table 3). For the model's fitness to increase, items with less than 0.6 and a MI of more than 10 should have been eliminated. However, they were kept because they were essential for the conceptual framework. Before arriving at the final model, 6 pairs of correlated errors were added to improve robustness. The resulting model was suitable for testing, as evidenced by the following model fit indexes: Chi Square: 641.705; df= 500; P< 0.001; CFI = 0.953; TLI: 0.948; RMSEA= 0.041 (90% CI 0.031 to 0.050).

Table 3: Result of Confirmatory Factor Analysis in phase 2 (n = 171)

Constructs	Items	Factor Loadings	AVE	CR
Perceived Susceptibility	PSus1	0.898	0.577	0.871
	PSus2	0.819		
	PSus3	0.739		
	PSus4	0.685		
	PSus5	0.683		
Perceived Severity	PSev1	0.680	0.597	0.881
	PSev2	0.675		
	PSev3	0.806		

	PSev4	0.867		
	PSev5	0.766		
	PBen1	0.752		
	PBen2	0.812		
	PBen3	0.740		
Perceived Benefit	PBen4	0.733	0.603	0.858
	PBar1	0.586		
	PBar2	0.670		
	PBar3	0.626		
	PBar4	0.886		
	PBar5	0.878		
Perceived Barrier	PBar6	0.770	0.572	0.888
	CUE1	0.646		
	CUE2	0.721		
	CUE3	0.727		
	CUE4	0.767		
Cue to action	CUE5	0.723	0.512	0.839
	EFF1	0.645		
	EFF2	0.694		
	EFF3	0.827		
	EFF4	0.859		
Self-efficacy	EFF5	0.744	0.574	0.869
	HBI1	0.654		
	HBI2	0.836		
Health Behavioural	HBI3	0.867		
Intention	HBI4	0.773	0.617	0.864

The AVE and CR values, which are listed in Table 3, were obtained after the structural model's fit was investigated to check if the items were loaded appropriately. The AVE readings were all over the cut-off value of 0.5, ranging from 0.512 to 0.617. The CR values were all over the cut-off value of 0.7, ranging from 0.839 to 0.888. All seven constructs featured sufficient convergent validity.

Discriminant validity was evaluated using Fornell-Larcker criteria by comparing the squared correlations and AVE scores for each of the pairwise constructs [34]. With the exception of Perceived Benefit < — > Cue to action, all paired constructs have shown established discriminant validity (Table 4).

Table 4: Result of Discriminant Validity in phase 2 (n = 171)

	Factor Correlation	Correlation Squared	Discriminant validity
Perceived Susceptibility<-->Perceived Severity	.312	0.097	Established
Perceived Susceptibility <-->Perceived Benefit	.260	0.068	Established
Perceived Susceptibility <-->Perceived Barrier	.204	0.042	Established
Perceived Susceptibility <-->Cue to Action	.172	0.030	Established
Perceived Susceptibility <-->Self-Efficacy	.236	0.056	Established
Perceived Severity <--> Perceived Benefit	.575	0.331	Established
Perceived Severity <--> Perceived Barrier	-.042	0.176	Established
Perceived Severity <--> Cue to Action	.238	0.057	Established
Perceived Severity <--> Self-Efficacy	.257	0.066	Established
Perceived Benefit <--> Perceived Barrier	-.085	0.007	Established
Perceived Benefit <--> Cue to Action	.826	0.682	Not Established
Perceived Benefit <--> Self-Efficacy	.349	0.122	Established
Perceived Barrier <--> Cue to Action	-.001	0.000	Established
Perceived Barrier <--> Self-Efficacy	-.157	0.025	Established
Cue to Action <--> Self-Efficacy	.600	0.360	Established

Internal consistency was deemed to be acceptable if the Cronbach's alpha coefficients were more than 0.7. According to reliability analyses, Cronbach α of the Perceived Susceptibility, Perceived Severity, Perceived Benefits, Perceived Barriers, Cues to Action, Self-Efficacy and Health Behavioural Intention were 0.83, 0.81, 0.86, 0.80, and 0.71, respectively in phase 1 and 0.87, 0.88, 0.86, 0.89, 0.84, 0.87 and 0.86, respectively in phase 2.

The detailed mean and standard deviation of each HBM component and Health Behavioural Intention to the socio-demographic characteristics are displayed in Table 5. The total mean score for each domain ranged from 14.03 to 21.32.

Table 5: Result of the minimum and maximum scores, total mean score and SD in phase 2 (n=171)

Constructs	N	Minimum	Maximum	Mean	SD
Perceived Susceptibility	171	5	25	14.03	4.04
Perceived Severity	171	5	25	16.81	4.44
Perceived Benefits	171	7	20	16.16	2.86
Perceived Barriers	171	6	30	21.32	4.99
Cue to Action	171	9	25	19.97	3.14
Self-Efficacy	171	9	25	18.04	3.87
Health Behavioural Intention	171	5	20	15.64	3.26

Discussion

At the time when the study was conducted, there were no published papers based on HBM to evaluate cancer health perception among at-risk smokers. Geography, ethnicity, national, social, and genetic-related factors contribute to the disproportionate burden of cancer. Sarawak's population vulnerability to NPC is among the highest in the world. Although genetic predisposition may be the most important risk factor leading to higher incidence of NPC in Sarawak, individual behavioural variables are a key driver of community health that must not be underestimated [35].

A total of 34 items in the questionnaire were formulated consistent with the HBM and divided into two sections: HBM scale for smokers' perception of NPC; and health behavioural intention to smoking cessation. Both sections are constructed with a 5-point Likert scale ranging from "1 = strongly disagree" to "5 = strongly agree". The first section consists of 30 items and is arranged into six subcategories, each representing the six constructs of the health belief model – perceived susceptibility (5 items), perceived seriousness (5 items), perceived benefits (4 items), perceived barriers (6 items), cues to action (5 items), and self-efficacy (5 items). The second section includes 4 items that predict health behavioural intention (See Table S1 and S2).

Based on the validity and reliability tests, including face and content validity, construct validity, and internal consistency, the findings of the current study indicate that the questionnaire has promising psychometric properties. Ten experts advised on content validity, and 7 items that did not reach the threshold of CVR based on Lawshe's Table and were judged superfluous were removed [25]. Face validity was examined in a pilot study with 10 smokers who fulfilled the eligibility requirements to ensure cultural acceptance and assess relevance and readability within the local community. Given a satisfactory Impact Score of each item, no item was eliminated in the face validity stage. In the main study, at-risk smokers from various ethnic groups participated, which was conducted in urban, suburban and rural regions of Miri (the northern region of Sarawak). The KMO test yielded a result of 0.697 (Phase 1) and 0.830 (Phase 2) while the Bartlett's test of sphericity obtained 1746.76 (Phase 1) and 3362.86 (Phase 2), both with p-value < 0.001, indicating that the sample size was adequate and the correlation between the items was sufficient for factor analysis.

Construct validity primarily concerns with the degree to which a concept measures what it claims to measure [36]. Similar to a previous study [37], a number of analyses were conducted to assess the construct validity, including EFA, CFA, as well as convergent and discriminant validity. The EFA demonstrated in phase 1 that the seven-factor structure accounted for 63.01 percent of the overall variance. The cut-off point for factor loading was fixed at 0.30. According to EFA suggestion, HBI2 item (I am trying to quit smoking to prevent nose and throat cancer at this time) should be grouped together with Perceived Benefit since the factor loading (0.537) is higher than when it is grouped with Health Behavioural Intention (0.307). This item requires immediate smoking cessation, which may spark inconsistent answers as some respondents may not be prepared to quit. The Cronbach's α coefficient value of 'Health Behavioural Intention' construct appeared to be satisfactory (Phase 1: 0.705; Phase 2: 0.861) and this item could potentially be essential on a larger scale. Thus, the researchers agreed not to delete this item.

In phase 2, CFA was performed to see if the seven-factor model derived by EFA could validate the association among the items based on the chosen framework. The loadings of all factors were more than 0.5. The goodness of fit was demonstrated to have an acceptable fit with the data, Chi Square = 641.705; df = 500; P< 0.001; CFI = 0.953; TLI: 0.948; RMSEA= 0.041 (90% CI 0.031 to 0.050). Following CFA, convergent and discriminant validity were tested in phase 2. The CR and AVE values for each component have to be higher than 0.7 and 0.5; respectively, which are fulfilled in the current study, suggesting an acceptable convergent validity [38]. Discriminant validity was evaluated between the

HBM constructs (excluding health behavioural intention construct). To establish acceptable discriminant validity, the factors' correlation coefficients with other factors must not be greater than each factor's AVE square root [34]. The current findings demonstrated established discriminant validity for all except for Perceived Benefit <-->Cue to Action. The explanation for this might be that the greater the perceived benefits of smoking cessation, the more likely smokers will look for cues to participate in such health-promoting behaviour, or vice versa. Future studies could delve deeper into the strength of the correlation, particularly between Perceived Benefit and Cue to Action. In terms of reliability, each construct for both phase 1 and 2 showed rationally acceptable Cronbach's α coefficient values as all of which were higher than 0.7, which demonstrates a high internal consistency [39].

This study has several limitations. Firstly, it is a cross-sectional study using convenience sampling, and thus, it is susceptible to recall and selection bias. However, to mitigate recall bias, participants were given ample time to consider their responses before answering the questions. This study was carried out with the cooperation of smokers from accessible areas in Sarawak. However, due to the widespread locations, the population sizes were difficult to determine, which contributed to the second limitation. The third limitation was the relatively small sample size for both EFA (n=100) and CFA (n=171). However, Kline (1994) indicated that for EFA, a sample size of 100 is sufficient, while Anderson & Gerbing (1988) suggested that CFA/SEM may be reliably examined with a minimum sample size of 100–150 [40, 41].

This study provides practical implications. With it being valid and reliable, the public health officials and researchers now have a reason to launch a larger population-based study on health behavioural intention to minimize NPC. Given that the current smoking rates in Malaysia remain high, this questionnaire can help in understanding and determining the factors that influence smokers' health decisions. Subsequently, cancer risk can be reduced by better prediction, a comprehensive tobacco control programme, policy creation, and health interventions.

Conclusions

In order to examine the variables affecting smoking cessation for cancer prevention, this study made an effort to develop a comprehensive HBM-based questionnaire. The results depict consistently satisfactory psychometric properties, confirming the validity and reliability. Considering that smoking is a major contributor to cancer, it is critical to address the health behavioural intention to uncover obstacles and implement improvements for a more successful intervention. The authors propose further studies to use the instruments for application in other smoking-related cancers in different susceptible populations and geographical locations.

Funding

Funding for the publication fee was received from RCSI & UCD Malaysia Campus.

Conflicts of Interest

The authors declare that there are no competing or potential conflicts of interest.

Author Contributions

Conceptualization: Kueh Martin, Fairuz Fadzilah Rahim, Abdul Rashid.

Formal analysis: Kueh Martin, Fairuz Fadzilah Rahim.

Investigation: Kueh Martin, Fairuz Fadzilah Rahim.

Methodology: Kueh Martin, Fairuz Fadzilah Rahim.

Project administration: Kueh Martin, Fairuz Fadzilah Rahim.

Resources: Kueh Martin, Fairuz Fadzilah Rahim.

Software: Kueh Martin.

Supervision: Fairuz Fadzilah Rahim, Abdul Rashid.

Validation: Kueh Martin.

Visualization: Kueh Martin.

Writing – original draft: Kueh Martin.

Writing – review & editing: Kueh Martin, Fairuz Fadzilah Rahim, Abdul Rashid.

Institutional Review Board Statement

The study was approved by Joint Penang Independent Ethics Committee on 13 August 2020 (Approval no. JPEC 20 0027).

Informed Consent Statement

Informed consent was obtained from all participants involved in the study.

Data Availability Statement

Extra data can be accessed via the Dryad data repository at http://datadryad.org/ with the doi: https://doi.org/10.5061/dryad.4tmpg4fbb

Acknowledgments

We express our gratitude to the hospital director in Miri and local estate managers for their helpful assistance and the participants of this study.

Word count 4236

References

1. Institute for Health Metrics and Evaluation (IHME). Findings from the Global Burden of Disease Study 2017. In: IHME, editor. Seattle, WA. 2018. Available from: http://www.healthdata.org/policy-report/findings-global-burden-disease-study-2017
2. World Health Organization. WHO Report on Cancer: Setting priorities, investing wisely and providing care for all. 2020. Available from: https://apps.who.int/iris/handle/10665/330745
3. Wellmann KF. [Smoking and Health. On the Report of the Advisory Committee to the Surgeon General of the Public Health Service]. Dtsch Med Wochenschr. 1964, 89, 1085-6.
4. Hecht SS. Research opportunities related to establishing standards for tobacco products under the Family Smoking Prevention and Tobacco Control Act. Nicotine & Tobacco Research. 2011, 14(1), 18-28. https://doi.org/10.1093/ntr/ntq216
5. Viswanath K, Herbst RS, Land SR, Leischow SJ, Shields PG. Tobacco and Cancer: An American Association for Cancer Research Policy Statement. Cancer Research. 2010, 70(9), 3419-30. https://doi.org/10.1158/0008-5472.can-10-1087
6. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021, 71(3), 209-49. https://doi.org/10.3322/caac.21660
7. Hashim D, Boffetta P, La Vecchia C, Rota M, Bertuccio P, Malvezzi M, et al. The global decrease in cancer mortality: trends and disparities. Ann Oncol. 2016, 27(5), 926-33. https://doi.org/10.1093/annonc/mdw027
8. Zavala VA, Bracci PM, Carethers JM, Carvajal-Carmona L, Coggins NB, Cruz-Correa MR, et al. Cancer health disparities in racial/ethnic minorities in the United States. Br J Cancer. 2021, 124(2), 315-32. https://doi.org/10.1038/s41416-020-01038-6
9. Lam KO, Lee AW, Choi CW, Sze HC, Zietman AL, Hopkins KI, et al. Global Pattern of Nasopharyngeal Cancer: Correlation of Outcome With Access to Radiation Therapy. Int J Radiat Oncol Biol Phys. 2016, 94(5), 1106-12. https://doi.org/10.1016/j.ijrobp.2015.11.047
10. Azizah A, Hashimah B, Nirmal K, Siti Zubaidah A, Puteri N, Nabihah A, et al. Malaysian National Cancer Registry Report 2012-2016 [Internet]. Malaysia: National Cancer Institute. . Putrajaya, Malaysia: National Cancer Institute, Ministry of Health, 2019. Available from: https://www.moh.gov.my/moh/resources/Penerbitan/Laporan/Umum/2012-2016%20(MNCRR)/MNCR_2012-2016_FINAL_(PUBLISHED_2019).pdf
11. Devi BCR, Pisani P, Tang TS, Parkin DM. High Incidence of Nasopharyngeal Carcinoma in Native People of Sarawak, Borneo Island. Cancer Epidemiology Biomarkers & Prevention. 2004, 13(3), 482-6.
12. Tsao SW, Yip YL, Tsang CM, Pang PS, Lau VM, Zhang G, et al. Etiological factors of nasopharyngeal carcinoma. Oral Oncol. 2014, 50(5), 330-8. https://doi.org/10.1016/j.oraloncology.2014.02.006
13. Linton RE, Daker M, Khoo ASB, Choo DCY, Viljoen M, Neilsen PM. Nasopharyngeal carcinoma among the Bidayuh of Sarawak, Malaysia: history and risk factors (Review). Oncol Lett. 2021, 22(1), 514. https://dx.doi.org/10.3892%2Fol.2021.12775
14. Guo S-S, Huang P-Y, Chen Q-Y, Liu H, Tang L-Q, Zhang L, et al. The impact of smoking on the clinical outcome of locoregionally advanced nasopharyngeal carcinoma after chemoradiotherapy. Radiation Oncology. 2014, 9(1), 246. https://doi.org/10.1186/s13014-014-0246-y
15. Xue WQ, Qin HD, Ruan HL, Shugart YY, Jia WH. Quantitative association of tobacco smoking with the risk of nasopharyngeal carcinoma: a comprehensive meta-analysis of studies conducted between 1979 and 2011. Am J Epidemiol. 2013, 178(3), 325-38. https://dx.doi.org/10.1093%2Faje%2Fkws479
16. Institute for Public Health. National Health and Morbidity Survey 2015 (NHMS 2015). Ministry of Health Malaysia Kuala Lumpur; 2015. Available from: https://www.moh.gov.my/moh/resources/nhmsreport2015vol2.pdf

17. Rahman M, Arif M, Razak MFA, Suhaili M, Tambi Z, Akoi C, et al. Factors associated with tobacco use among the adult population in Sarawak, Malaysia: A cross sectional study. Epidemiology Biostatistics and Public Health. 2015, 12, 1-9. https://doi.org/10.2427/10292
18. Champion VL, Skinner CS. The health belief model. Health behavior and health education: Theory, research, and practice, 4th ed. San Francisco, CA, US: Jossey-Bass, 2008. p. 45-65.
19. Maiman LA, Becker MH. The Health Belief Model: Origins and Correlates in Psychological Theory. Health Education Monographs. 1974, 2(4), 336-53. https://doi.org/10.1177/109019817400200404
20. Odedina FT, Dagne G, Pressey S, Odedina O, Emanuel F, Scrivens J, et al. Prostate cancer health and cultural beliefs of black men: The Florida Prostate Cancer Disparity Project. Infectious Agents and Cancer. 2011, 6(2), S10. https://dx.doi.org/10.1186%2F1750-9378-6-S2-S10
21. Solomon K, Tamire M, Kaba M. Predictors of cervical cancer screening practice among HIV positive women attending adult anti-retroviral treatment clinics in Bishoftu town, Ethiopia: the application of a health belief model. BMC Cancer. 2019, 19(1), 989. https://doi.org/10.1186/s12885-019-6171-6
22. Tavafian SS, Hasani L, Aghamolaei T, Zare S, Gregory D. Prediction of breast self-examination in a sample of Iranian women: an application of the Health Belief Model. BMC Women's Health. 2009, 9(1), 37. https://doi.org/10.1186/1472-6874-9-37
23. LaMorte., Wayne. W. The Health Belief Model: Behavioral Change Models 2019. Available from: http://sphweb.bumc.bu.edu/otlt/MPH-Modules/SB/BehavioralChangeTheories/BehavioralChangeTheories2.html
24. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine. 2000, 25(24), 3186-91. https://doi.org/10.1097/00007632-200012150-00014
25. Lawshe CH. A quantitative approach to content validity. Personnel psychology. 1975, 28(4), 563-75. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x
26. Samet JM, Wipfli H, Platz EA, Bhavsar N. A Dictionary of Epidemiology, Fifth Edition: Edited by Miquel Porta. American Journal of Epidemiology. 2009, 170(11), 1449-51. https://doi.org/10.1093/aje/kwp322
27. Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. Frontiers in public health, 6, 149. https://doi.org/10.3389/fpubh.2018.00149
28. Lacasse Y, Godbout C, Sériès F. Health-related quality of life in obstructive sleep apnoea. Eur Respir J. 2002, 19(3), 499-503. https://doi.org/10.1183/09031936.02.00216902
29. Suhr DD. Exploratory or confirmatory factor analysis? 2006. Available from: https://support.sas.com/resources/papers/proceedings/proceedings/sugi31/200-31.pdf
30. [dataset] Kueh M, Rahim FF, Rashid A. Data from: Development and validation of the health behavioural intention on smoking cessation to prevent nasopharyngeal cancer in Sarawak, Malaysia based on the Health Belief Model: a cross-sectional study. Dryad Digital Repository, November 18, 2021. https://doi.org/10.5061/dryad.4tmpg4fbb
31. Munro BH. Statistical methods for health care research: lippincott williams & wilkins; 2005.
32. Xiang B, Wong HM, Cao W, Perfecto AP, McGrath CPJ. Development and validation of the Oral health behavior questionnaire for adolescents based on the health belief model (OHBQAHBM). BMC Public Health. 2020, 20(1), 701. https://doi.org/10.1186/s12889-020-08851-x
33. Brown TA. Confirmatory factor analysis for applied research. New York, NY, US: The Guilford Press; 2006.
34. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. Journal of marketing research. 1981, 18(1), 39-50. https://doi.org/10.2307/3151312
35. McGinnis JM, Williams-Russo P, Knickman JR. The case for more active policy attention to health promotion. Health affairs. 2002, 21(2), 78-93. https://doi.org/10.1377/hlthaff.21.2.78

36. Cohen RJ, Swerdlik ME, Phillips SM. Psychological testing and assessment: An introduction to tests and measurement: Mayfield Publishing Co; 1996.
37. Shamsalinia A, Moradi M, Farahani MA, Masoudi R, Ghadimi R, Rad RE, et al. Designing and psychometric evaluation of disease-related fear scale (D-RFS) in adults with epilepsy: A sequential exploratory mixed methods design. Epilepsy & Behavior. 2020, 110, 107169. https://doi.org/10.1016/j.yebeh.2020.107169
38. Schreiber JB, Nora A, Stage FK, Barlow EA, King J. Reporting structural equation modeling and confirmatory factor analysis results: A review. The Journal of educational research. 2006, 99(6), 323-38. https://doi.org/10.3200/JOER.99.6.323-338
39. Kline P. The handbook of psychological testing: Psychology Press, 2000. https://doi.org/10.4324/9781315812274
40. Kline P. An Easy Guide to Factor Analysis (1st ed.). 1st edition ed. London. 1994, p. 208. https://doi.org/10.4324/9781315788135
41. Anderson JC, Gerbing DW. Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. Psychological Bulletin. 1988, 103, 411-423. http://dx.doi.org/10.1037/0033-2909.103.3.411

Figure 1. Flow diagram of the development and validation of the Health Belief Model questionnaire to promote smoking cessation for nasopharyngeal cancer prevention.

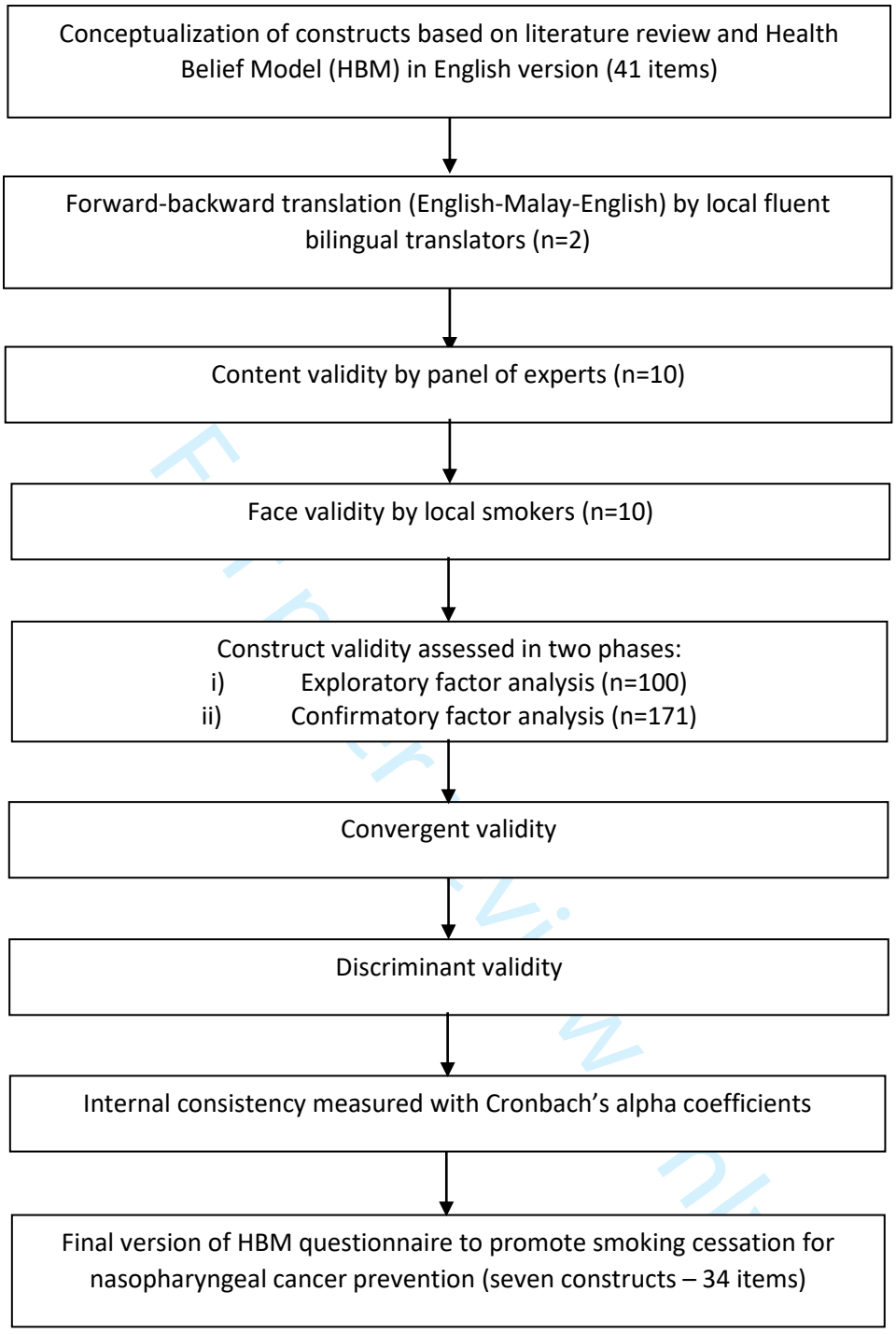

Figure 1. Flow diagram of the development and validation of the Health Belief Model questionnaire to promote smoking cessation for nasopharyngeal cancer prevention.

Table S1: Questionnaire to promote smoking cessation for nasopharyngeal cancer prevention (English version).

Section 1: Perceptions towards nose and throat cancer

Please choose the best answer options to indicate your level of agreement.

Constructs of Health Belief Model	Questions	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Perceived susceptibility	1. I feel that I will get nose and throat cancer in the future.					
	2. My chances of getting nose and throat cancer are high.					
	3. I cannot avoid myself from getting nose and throat cancer.					
	4. I am worried about getting nose and throat cancer due to family history.					
	5. My smoking habit makes me more likely than average to get nose and throat cancer.					
Perceived seriousness	6. The thought of nose and throat cancer scares me.					
	7. If I had nose and throat cancer, the cost of treatment can be a financial strain.					
	8. Nose and throat cancer would threaten a relationship with my family.					

	9. If I had nose and throat cancer my whole life would change.					
	10. If I developed nose and throat cancer, I would not live long.					
Perceived benefits	11. Quit smoking decreases my chance of getting nose and throat cancer.					
	12. Quit smoking decreases my chance of dying from nose and throat cancer.					
	13. Quit smoking can improve my health.					
	14. I feel less anxious about nose and throat cancer if I quit smoking.					
Perceived barriers	15. It is difficult to quit tobacco smoking (e.g., peer pressure).					
	16. I feel anxious without smoking.					
	17. Tobacco smoking relieves my stress.					
	18. I experience headache without smoking.					
	19. I experience excessive salivation without smoking.					
	20. I think getting nose and throat cancer is destiny and quitting smoking will not change it.					
Cues to action	21. I will quit smoking if I have social support.					

	22. I will quit smoking if there are information sources that reminds me. (Examples of sources include: the internet, newspapers, radio and TV).					
	23. I will quit smoking if I have the will to change.					
	24. I will quit smoking if I know the diseases related to smoking.					
	25. I will quit smoking if there are health professionals to assist me.					
Self-efficacy	26. I can refuse to smoke when I am thinking about difficult problem.					
	27. I can refuse the urge to smoke.					
	28. I can refuse to smoke when I see someone else smoking.					
	29. I can refuse to smoke when offered by my friends/ family.					
	30. I can refuse from buying cigarettes when I have extra pocket money.					

Section 2: Health Behavioural Intention on smoking cessation

Please choose the best answer options to indicate your level of agreement.

Statements	Strongly disagree	Disagree	Neutral	Agree	Strongly agree

1. I would like to lead a healthier lifestyle to prevent nose and throat cancer.					
2. I am trying to quit smoking to prevent nose and throat cancer at this time.					
3. I plan to quit smoking to prevent nose and throat cancer within six months.					
4. I would like to quit smoking to prevent nose and throat cancer but have never really tried.					

Table S2: Questionnaire to promote smoking cessation for nasopharyngeal cancer prevention (Malay version- *Borang kajiselidik untuk menggalakkan niat tingkah laku kesihatan berkenaan berhenti merokok bagi pencegahan kanser pangkal hidung*)

Bahagian 1: Persepsi terhadap kanser hidung dan tekak.

Sila tandakan pilihan jawapan terbaik untuk menunjukkan tahap persetujuan anda.

Domain Model Kepercayaan Kesihatan	Soalan	Sangat tidak setuju	Tidak setuju	Kurang bersetuju	Setuju	Sangat setuju
Kepercayaan kepada tiada daya tahan	1. Saya merasakan bahawa saya akan mendapat kanser pangkal hidung pada masa akan datang.					
	2. Peluang saya mendapat kanser pangkal hidung adalah tinggi.					
	3. Saya tidak dapat mengelakkan diri daripada mendapat kanser pangkal hidung.					
	4. Saya bimbang terkena kanser pangkal hidung kerana ahli keluarga saya menghadapi masalah ini.					
	5. Tabiat merokok saya menyebabkan saya lebih berisiko menghidap kanser pangkal hidung.					
Kepercayaan kepada bahaya	6. Risiko dan kesan-kesan kanser pangkal hidung menakutkan saya.					
	7. Jika saya mempunyai kanser pangkal hidung, kos					

	rawatan akan menjadi beban kewangan keluarga saya.					
	8. Kanser pangkal hidung akan merosakkan hubungan antara ahli keluarga saya.					
	9. Jika saya mempunyai kanser pangkal hidung, kehidupan saya akan berubah.					
	10. Jika saya mempunyai kanser pangkal hidung, saya tidak akan hidup lama.					
Kepercayaan kepada manfaat	11. Berhenti merokok akan mengurangkan peluang saya mendapat kanser pangkal hidung.					
	12. Berhenti merokok akan mengurangkan risiko kematian disebabkan kanser pangkal hidung.					
	13. Berhenti merokok amat penting untuk meningkatkan kesihatan saya.					
	14. Berhenti merokok akan mengurangkan kerisauan saya terhadap kanser pangkal hidung.					
Kepercayaan kepada halangan	15. Saya berasa sukar untuk menghentikan tabiat merokok atas					

	sebab tekanan rakan sebaya dan lain-lain.					
	16. Saya akan berasa bimbang akan sesuatu tanpa merokok.					
	17. Merokok menghilangkan tekanan saya.					
	18. Jika saya tidak merokok, saya akan mengalami sakit kepala.					
	19. Jika saya tidak merokok, saya akan mengalami air liur berlebihan.					
	20. Saya rasa mendapat kanser pangkal hidung adalah takdir dan berhenti merokok tidak akan mengubahnya.					
Isyarat untuk bertindak	21. Jika saya mempunyai sokongan sosial, saya akan berhenti merokok.					
	22. Jika beberapa sumber maklumat mengingatkan saya, saya akan berhenti merokok. (Contoh sumber maklumat termasuk: internet, surat khabar, radio dan TV).					
	23. Saya akan berhenti merokok sekiranya saya mempunyai kemahuan untuk berubah.					

	24. Menyedari bahaya merokok akan membantu saya berhenti merokok.					
	25. Saya akan berhenti merokok sekiranya ahli profesional kesihatan membantu saya.					
Keberkesanan diri	26. Saya boleh menolak merokok apabila saya ingin memikirkan masalah yang sukar.					
	27. Saya boleh menolak merokok apabila saya mempunyai keinginan untuk merokok.					
	28. Saya boleh menolak merokok apabila saya melihat orang lain merokok.					
	29. Saya boleh menolak merokok apabila kawan/ keluarga saya mengajak saya untuk merokok.					
	30. Saya boleh menolak membeli rokok apabila saya mempunyai lebihan wang saku.					

Bahagian 2: Tingkah laku untuk berhenti merokok demi kesihatan

Sila tandakan pilihan jawapan terbaik untuk menunjukkan tahap persetujuan anda.

Soalan	Sangat tidak bersetuju	Tidak setuju	Kurang bersetuju	Setuju	Sangat setuju

1. Saya akan menjalani gaya hidup yang lebih sihat untuk mencegah kanser pangkal hidung.					
2. Saya sedang berusaha untuk berhenti merokok untuk mengurangkan risiko kanser pangkal hidung pada masa ini.					
3. Saya merancang untuk berhenti merokok untuk mengurangkan risiko kanser pangkal hidung dalam masa enam bulan.					
4. Saya ingin berhenti merokok untuk mengurangkan risiko kanser pangkal hidung tetapi tidak pernah mencuba.					

Reporting checklist for cross sectional study.

Based on the STROBE cross sectional guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the STROBE cross sectionalreporting guidelines, and cite them as:

von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies.

		Reporting Item	Page Number
Title and abstract			
Title	#1a	Indicate the study's design with a commonly used term in the title or the abstract	3

Abstract	#1b	Provide in the abstract an informative and balanced summary of what was done and what was found	3
Introduction			
Background / rationale	#2	Explain the scientific background and rationale for the investigation being reported	4-5
Objectives	#3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	#4	Present key elements of study design early in the paper	6-7
Setting	#5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	7
Eligibility criteria	#6a	Give the eligibility criteria, and the sources and methods of selection of participants.	7
	#7	Clearly define all outcomes, exposures, predictors, potential	7

		confounders, and effect modifiers. Give diagnostic criteria, if applicable	
Data sources / measurement	#8	For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Give information separately for for exposed and unexposed groups if applicable.	8-9
Bias	#9	Describe any efforts to address potential sources of bias	15. We involved experts and participants of different background and ethnic groups in hopes to reduce biasness. We are cognizant that convenience sampling was approached, which gives rise to selection and recall bias. We highlighted as the first limitation in the manuscript.
Study size	#10	Explain how the study size was arrived at	15
Quantitative variables	#11	Explain how quantitative variables were handled in the analyses. If	7-8

		applicable, describe which groupings were chosen, and why	
Statistical methods	#12a	Describe all statistical methods, including those used to control for confounding	8-9
Statistical methods	#12b	Describe any methods used to examine subgroups and interactions	8-9
Statistical methods	#12c	Explain how missing data were addressed	9
Statistical methods	#12d	If applicable, describe analytical methods taking account of sampling strategy	9
Statistical methods	#12e	Describe any sensitivity analyses	N/A
Results			
Participants	#13a	Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed. Give information separately for for	N/A. Do not have a list of numbers potentially eligible.

		exposed and unexposed groups if applicable.	
Participants	#13b	Give reasons for non-participation at each stage	N/A
Participants	#13c	Consider use of a flow diagram	8
Descriptive data	#14a	Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Give information separately for exposed and unexposed groups if applicable.	9-10
Descriptive data	#14b	Indicate number of participants with missing data for each variable of interest	N/A. No missing data was collected
Outcome data	#15	Report numbers of outcome events or summary measures. Give information separately for exposed and unexposed groups if applicable.	N/A. This study focuses on validating the questionnaire with a series of analysis.
Main results	#16a	Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were	N/A. This study focuses on validating the questionnaire with a series of analysis.

		adjusted for and why they were included	
Main results	#16b	Report category boundaries when continuous variables were categorized	N/A. This study focuses on validating the questionnaire with a series of analysis.
Main results	#16c	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A. This study focuses on validating the questionnaire with a series of analysis.
Other analyses	#17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	9-13
Discussion			
Key results	#18	Summarise key results with reference to study objectives	14-15
Limitations	#19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.	15
Interpretation	#20	Give a cautious overall interpretation considering objectives, limitations, multiplicity of analyses, results from	14

		similar studies, and other relevant evidence.	
Generalisability	#21	Discuss the generalisability (external validity) of the study results	14-15
Other Information			
Funding	#22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	15

Notes:

- 9: p. 15: We involved experts and participants of different background and ethnic groups in hopes to reduce biasness. We are cognizant that convenience sampling was approached, which gives rise to selection and recall bias. We highlighted as the first limitation in the manuscript.
- 13a: N/A. Do not have a list of numbers potentially eligible.
- 14b: N/A. No missing data was collected
- 15: N/A. This study focuses on validating the questionnaire with a series of analysis.
- 16a: N/A. This study focuses on validating the questionnaire with a series of analysis.
- 16b: N/A. This study focuses on validating the questionnaire with a series of analysis.
- 16c: N/A. This study focuses on validating the questionnaire with a series of analysis. The STROBE checklist is distributed under the terms of the Creative Commons Attribution License