CHAPTER 3 : RESEARCH METHODOLOGY

3.1 Introduction

This chapter gives a detailed explanation of the methods used to collect the data and how the findings were derived. Section 3.2 describes the research design and rationale of the design used in this study. Research procedures for systematic review (phase one), quantitative research (phase two) and the qualitative study (phase three) are also defined accordingly. Figure 3.1 shows a visual research flow for the three-phase architecture of the systematic reciew and explanatory sequential mixed-method approach. Section 3.3 describes the methods used in the systematic review phase, while section 3.4 describes the methods used in the quantitative phase, and lastly section 3.5 describes the methods used in the qualitative phase.

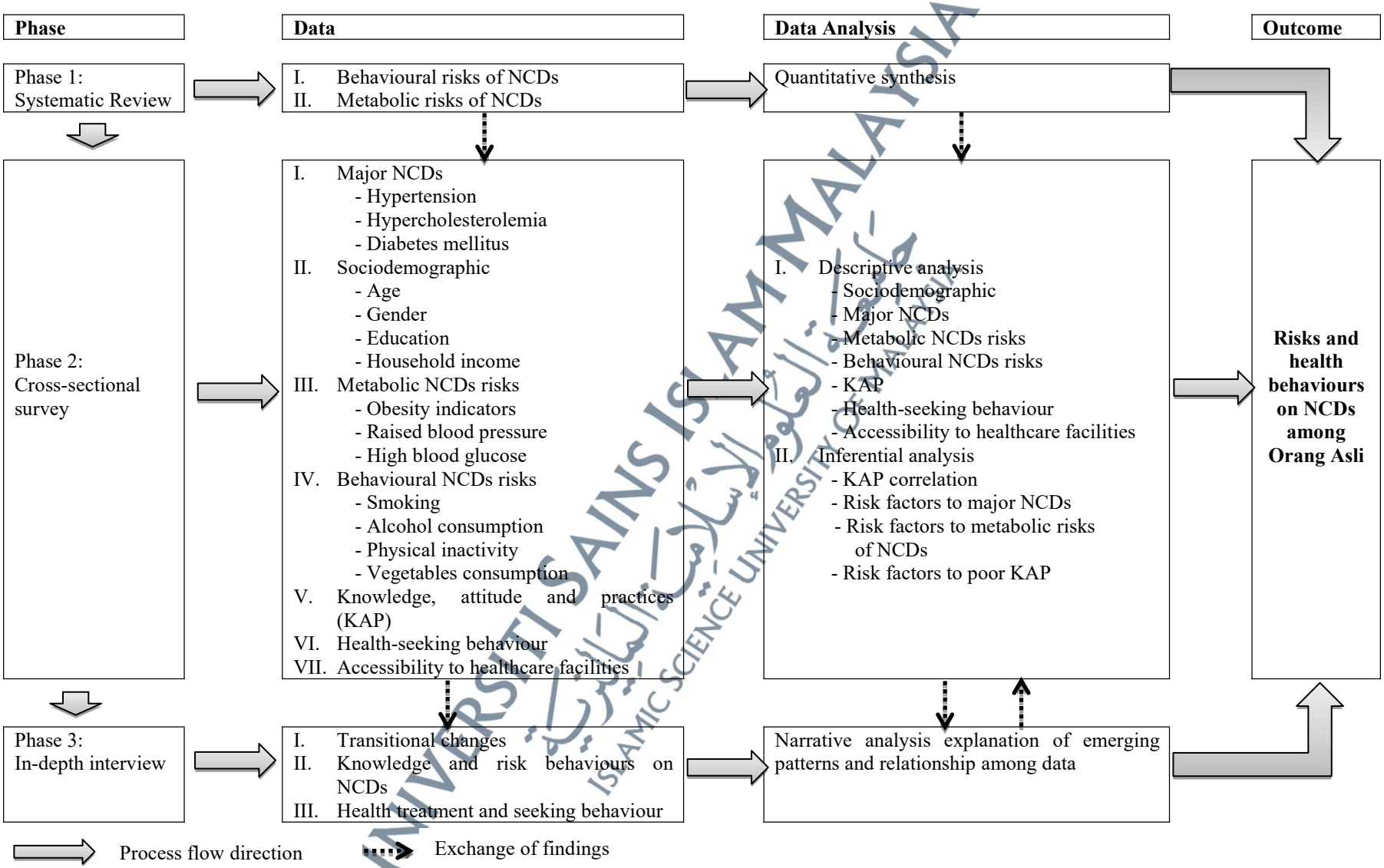


Figure 3.1: Research Flow for Systematic Review and An Explanatory Sequential Mixed-Methods Design

3.2 Research Design

Research design refers to the use of the most effective way of achieving the study's goal at the minimum cost. Systematic review research involves a detailed and comprehensive plan to recognise, appraise, and synthesise all related studies on a topic. The search strategy is derived a priori to reduce bias (Uman, 2011). Welch, Boyer & Chamberlain (2015) reported that results from the systematic review can provide valuable evidence to support the decision-making process in improving Indigenous people's health. It is a research process that is vital when conducting a study on community health, as evidenced by previous reviews that successfully influenced the practice and policy (Kinchin et al., 2017).

The systematic review in this study was to investigate thoroughly the areas relevant to the study research questions and to identify the research gaps in order to guide the directions of this thesis. As recommended by Hayman, Reid & King (2015) and Kinchin et al. (2017), the planning of community awareness and intervention programmes as well as the provision of guidance for policy implementations must include the following prerequisites:

a) directly relating to the topics under research

b) immersed in the medical and health knowledge surrounding the topics

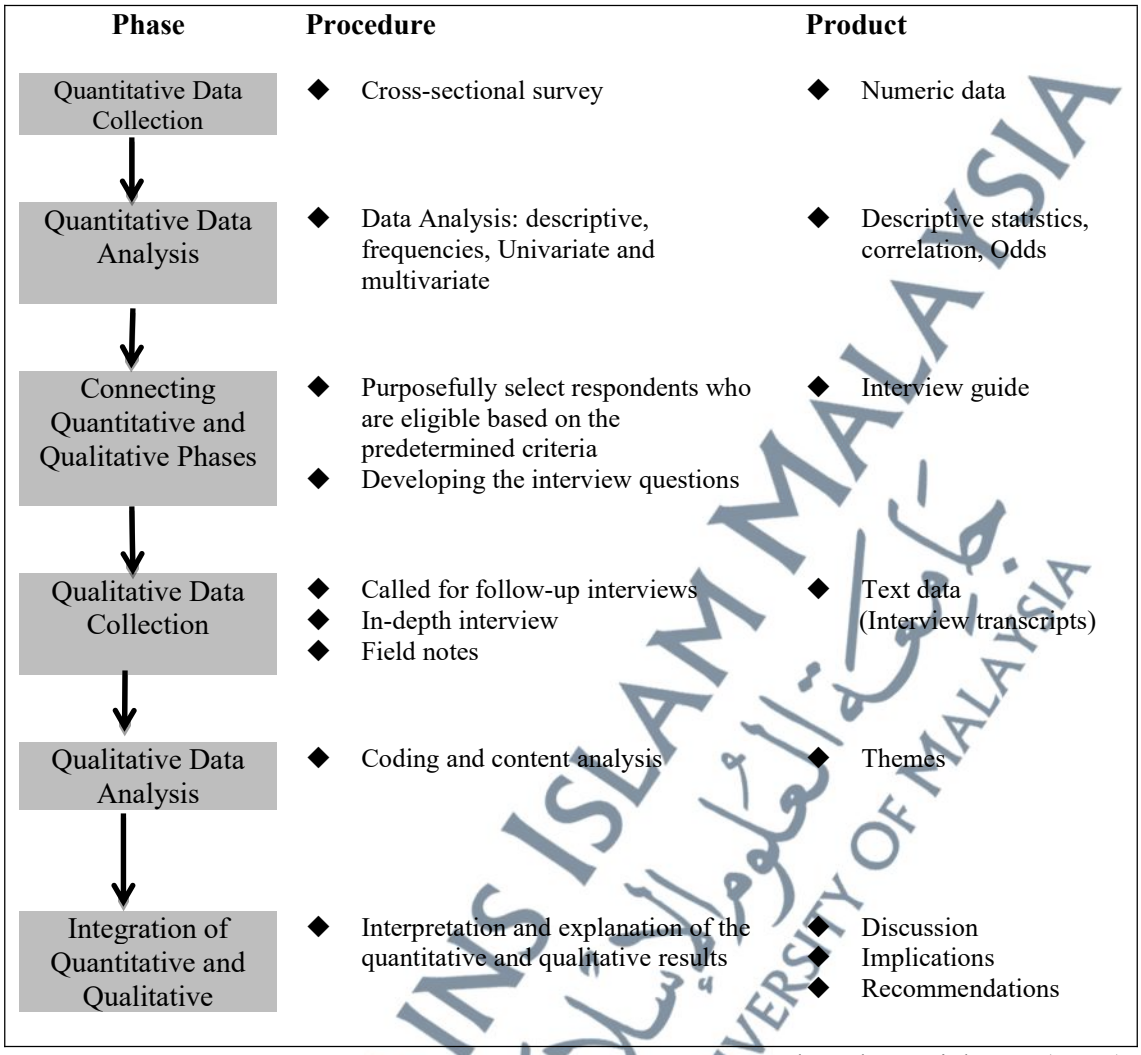
c) assessing conflicting research to determine the best study plans

Then, a cross-sectional study mixed-methods design for the data collection applied. By using this method, data can be collected at a specific time. There are different types of design categories, including explanatory, exploratory, parallel, and nested designs (Halcomb & Hickman, 2015). The choice of design depends on what is the most appropriate way to address the research questions and to produce reliable results.

Mixed-methods design is a research approach whereby researchers collect and analyse both quantitative and qualitative data within the same study (Creswell, 2014). The mixed-methods design can be of different sequence combinations and the methods used can also vary in emphasis. It confers the potential strengths of both qualitative and quantitative methods by exploring multiple perspectives, i.e. providers, policymakers, recipients, to determine the relationships between the research questions. A mixed-methods approach is able to provide a valuable source of data for community health research.

In this study, a quantitative followed by an explanatory qualitative study was the chosen sequence of mixed-methods. The data for this research were obtained from both quantitative and qualitative phases, collected and analysed separately then combined (Creswell, 2014). The method provided a strong inference and apart from complementing the strength of each methodology to provide more meaningful and in-depth data (Dahlan, Nicol & Maciver, 2010). The mixed-methods sequential explanatory model is presented in Figure 3.2.

Equal stress and priority were given to both the quantitative and qualitative phases. As discussed in Johnson, Onwuegbuzie & Turner (2007), Creswell (2014), and Shorten & Smith (2017), specific criteria unique to mixed-methods were given additional consideration. The following are the issues that needed to be addressed in the methodology:



Source: Halcomb & Hickman (2015)

Figure 3.2: Mixed-methods Sequential Explanatory Design Procedures

a) As the study design aimed to achieve an equal balance, the weighting of the quantitative to qualitative components is important.

b) Being a sequential design, recruitment of respondents for Phase Two would be dependent upon the assessment of the respondents in Phase One.

c) The reasons for the selection of the respondents of both phases were explicitly stated in the proposal and respondent information sheets. This assured the study rigour as it was carefully designed to gather the data for both the quantitative

Phase One and the qualitative Phase Two studies from the same set of respondents.

d) Several procedures further improved the validity of the research. In addition to providing further description of the quantitative results, meaningful information was also extracted from the dialogue in Phase Two. The quantitative outcomes were also compared and contrasted with the qualitative findings as part of the triangulation design method to verify or refute the study results. In short, all aspects of the study process were rigorous and systematic based on the mixed-methods methodological standards.

3.2.1 Systematic Review Overview

A systematic review conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses guideline (The PRISMA Statement) by Moher et al., (2009). It focuses explicitly on the modifiable behavioural factors and metabolic risks that may predict the development and advancement of major types of NCDs. The guiding questions for systematic review was:

a) 'What is the prevalence of behavioural risk factors of NCDs among the adults Orang Asli in Malaysia?'

b) 'What is the prevalence of metabolic risk factors of NCDs among the adults Orang Asli in Malaysia?'

3.2.2 Quantitative Overview

From the conceptual framework, the variables included in the quantitative study, including:

i. Sociodemographic factors: age, gender, education level, economic status

ii. Disease history: Major NCDs of hypertension, hypercholesterolemia and Diabetes mellitus

iii. Metabolic risk factors of NCDs: raised blood pressure, high blood glucose, obesity indicators

iv. Behavioural risk factors of NCDs: smoking, alcohol consumption, physical inactivity and inadequate vegetable intake

v. Knowledge, attitude and practice towards NCDs

vi. Health-seeking behaviour

The quantitative methods used to investigate the exposures of interest and to screen the respondents included:

i. Standardised face to face interview questionnaire on respondents' background: disease history, metabolic and behavioural risk factors of NCDs, KAP and health-seeking behaviour

ii. Body examination: height, weight, waist circumference, body fat composition measurement, blood pressure, blood glucose

From the data collected, trends and significant factors associated with NCDs among the respondents were determined.

3.2.3 Qualitative Overview

After the quantitative data collection, the second phase of the qualitative study was conducted. By carrying out this qualitative research, the respondents could express their views and perspectives more openly and liberally. These results were less likely to be achieved through quantitative analysis that reported only numbers and figures. The theme used in this qualitative data collection was in parallel with the

objectives of the quantitative study. The qualitative data collection utilised semi-structured in-depth interviews with consenting respondents.

3.3 Methods of the Systematic Review

3.3.1 Protocol and registration

The systematic review rational and method presented and approved at the candidature qualifying test at the Faculty of Medicines and Health Sciences, USIM. Systematic review protocol was not registered with any databases for publication of the systematic review or meta-analysis protocols as registration was not required by USIM institution.

3.3.2 Eligibility criteria

The inclusion and exclusion criteria for the retrieved articles were formulated with the principal supervisor committee Associate Prof. Dr. Mohd Dzulkhairi bin Mohd Rani. A discussion was also held with Dr. Khairun Nain bin Nor Aripin, a medical lecturer who works at the Faculty of Medicine and Health Sciences, Universiti Sains Islam Malaysia. His research areas are mainly in systematic reviews and evidence-based medicine.

The criteria developed were tested on two batches using the titles, abstracts, and the full-texts of twenty articles regarding their relevance, understandability, and practicality by the author and Dr. Nadia binti Mohd Effendy, a member on the supervisor committee. In the systematic review, having more than one researcher to develop and test the criteria can make the article retrieval process more objective (Rousseau, Manning & Denyer, 2008).

3.3.2.1 Behavioural Risk Factors of NCDs

The selection criteria for behavioural risk factors of NCDs are summarised in Table 3.1. The review included observational studies that estimated the prevalence of behavioural risk factors for NCDs among Orang Asli. The respondents were adult Orang Asli as defined by the original study authors. However, studies conducted on indigenous populations outside Peninsular Malaysia and Orang Asli categorised under other ethnic groups will be excluded.

The outcome measure was the prevalence of the four behavioural NCDs risk factors as identified by IPH (2015b) and WHO (2017), namely tobacco use, harmful use of alcohol, physical inactivity, and unhealthy diet. The prevalence should be measured by one (or more) of the following methods of routine data sources, i.e. medical notes, prescription lists, physical examinations, diagnostic tests, surveys, interviews, or self-reported answers.

Table 3.1: Inclusion and Exclusion Criteria for Behavioural Risk Factors of NCDs

	Inclusion	Exclusion
Respondents	● Adult as defined by author ● Orang Asli as defined by the author ● Male or Female ● Any length of time in data collection ● Whole general Orang Asli population or a defined subtribe	● Orang Asli categorised together with other ethnic groups in Sabah or Sarawak
Disease	● Behavioural Risk for NCDs (IPH, 2015b; WHO, 2017) -Tobacco smoke -Alcohol use -Low physical activity -Dietary risks	● Mental and substance use ● Suicide and self-harm ● Communicable disease ● Others NCDs risk factor ● NCDs
Outcome	● Prevalence measured from routine data sources such as hospital databases, medical notes, prescription lists, physical examinations, diagnostic tests, survey, or self-reported	● No number or count data
Study	● Prevalence studies using cross-sectional studies, cohort studies, or surveys. ● Conducted in Peninsular Malaysia ● Limited to publication until 31 December 2018 ● Limited to English and Malay language	● Qualitative ● Policy ● Opinion ● Case-studies ● Case-series studies

3.3.2.2 Metabolic Risk Factors of NCDs

Eligible studies included observational studies conducted among any Orang Asli population that stratified the respondents based on the presence or absence of the metabolic syndrome. The studies must report the outcomes as count data or prevalence and published in either English or Malay language. Studies investigating more than one metabolic syndrome or more than one definition of the metabolic syndrome were also eligible for inclusion. Studies that did not meet these criteria were excluded. The summaries of selection criteria are shown in Table 3.2.

Table 3.2: Inclusion and Exclusion Criteria for Metabolic Risk Factors of NCDs

	Inclusion	Exclusion
Respondents	● Adult as defined by author ● Orang Asli as defined by the author ● Male or Female ● Any length of time in data collection ● Whole general Orang Asli population or a defined sub-tribe	● Orang Asli categorised together with other ethnic groups in Sabah or Sarawak
Disease	● Metabolic syndrome as identified by the author ● Any cluster of metabolic risk factors of NCDs	● Mental and substance use ● Suicide and self-harm ● Communicable disease
Outcome	● Prevalence measured from routine data sources such as; hospital databases, medical notes, prescription lists, physical examinations, diagnostic tests	● No number or count data
Study	● Prevalence studies using cross-sectional studies, cohort studies, or surveys. ● Conducted in Peninsular Malaysia ● Limited to publication until 12 September 2019 ● Limited to English and Malay language	● Qualitative ● Policy ● Opinion ● Case-studies ● Case-series studies

3.3.3 Information sources

We searched four of the databases listed below:

i. CINAHL®Complete@EBSCOhost

ii. Medline®Complete @EBSCOhost

iii. PubMed®

iv. Scopus®

Four databases were selected to ensure all the relevant papers were included (Crossan & Apaydin, 2010). All four databases are well-established, multi-disciplinary research platforms. They include an extensive collection of peer-reviewed journals that are being regularly kept up to date. Before being accepted for publication in scholarly journals, all the articles would undergo a rigorous peer-review process. According to the McKiernan et al. (2019), peer review is defined as a board of scholarly reviewers in the subject area of the journal and they review the materials they intend to publish to ensure the quality of the research and adherence to editorial standards of the journal before accepting the articles for publication. This process ensures that the articles published in the journal are of higher quality.

According to Greenhalgh & Peacock (2005), a systematic review of complex evidence cannot solely rely on predefined, protocol-driven search strategies, regardless of the number of databases searched. Thus, it is necessary to conduct several search strategies to increase the effectiveness and efficiency of search methods.

Additionally, after an extensive electronic search, personal knowledge or personal contact and reference tracking were also conducted. In this study, the authors personally contacted the Institute for Public Health (IPH) Malaysia, the government organisation that conducted the Malaysian NHMS 2015 to obtain any grey or unpublished studies or data. Appendix 1 shows the request letter for data addressed to IPH Malaysia.

The snowballing method was also used to search for any additional citation related to the research question from the reference list. Snowballing methods are

especially useful in identifying sources from references of references and electronic citation tracking (Greenhalgh & Peacock, 2005).

3.3.4 Search strategy

Using the Population, Intervention, Comparison and Outcomes (PICO's) approach (Moher et al., 2009), the literature search based on four concepts: the population (adults Orang Asli), the type of measure (prevalence), outcome (behavioural or metabolic risks of NCDs) and the geographic setting (Malaysia). The list the search terms used was derived from the purpose statement after identifying the concepts of interest. The search terms were developed in collaboration with the committee supervisors and an expert in the systematic review field, Dr. Khairun Nain bin Nor Aripin. References were made using previously conducted systematic reviews of similar topics. No language limitations with regard to English and Indonesian publications

3.3.4.1 Behavioural Risk Factors of NCDs

The full search strategy for PubMed® is available in Appendix 2. These terms were searched in all the databases for full texts (Howard, Arnsten & Gourevitch, 2004; Irwin, 2004; Willi et al., 2007; Forouzanfar et al., 2016):

a) (nutrition OR diet OR nutrition assessment OR nutrition surveys) OR (exercise OR physical inactivity OR physical activity OR motor activity OR sports) OR (smoking OR smoking cessation OR smoke inhalation injury OR tobacco smokeless OR tobacco use cessation OR tobacco use disorder OR tobacco OR nicotine OR nicotine dependence OR tobacco dependence OR smoking dependence OR cigarette) OR (ethanol OR alcoholism OR alcoholic beverages OR alcohol-related disorders OR alcohol drinking)

AND

b) Orang Asli OR indigenous OR aboriginal OR native OR proto-malay OR negrito OR senoi

AND

c) Malaysia

3.3.4.2 Metabolic Risk Factors of NCDs

The full search strategy for PubMed® is available in Appendix 3. These terms were searched for the full texts on metabolic syndrome literature (Mottillo et al., 2010; Rodríguez-Monforte et al., 2017).

a) metabolic syndrome OR syndrome X OR metabolic syndrome X

AND

b) Orang Asli OR indigenous OR native OR aboriginal OR original OR proto-Malay OR Senoi OR negrito

AND

c) Malaysia

3.3.5 Study Selection

Duplicate studies from the different electronic databases, which emerge after using the above search strategy, removed through the EndNote X9 (Clarivate Analytics; Philadelphia, Pennsylvania). By referring to the Preferred Reporting Items for Systematic reviews and Meta-Analyses Group (PRISMA), the relevance of the retrieved articles were assessed in a four-step process. They were evaluated based on the in- and exclusion criteria (Moher et al., 2009). After the duplicate citations were removed, the study eligibility was assessed based on the pre-specified inclusion criteria.

Two reviewers consisting of the author and committee member, Dr. Nadia bin Mohd Effendy independently screened each title and abstract and categorised the study as “include”, “exclude”, or “unsure”. The list of titles and abstracts was compared and re-categorised if there was disagreement between the two reviewers. Where there was an agreement to exclude the citation, the citation would be excluded at this stage. Articles that were agreed by both to include or labelled as unsure were reviewed by reading the full text. If the study remained unsure after reading the full text, it would be reviewed by a third reviewer, Dr. Khairun Nain bin Nor Aripin. The reasons for exclusion were documented at the stage of the screening the full-text.

3.3.6 Data Collection Process

The complete list of articles that met the inclusion criteria was then selected and documented. Each selected article was also entered into a reference management database using Endnote X9. Data extracted using a standardised extraction form in Microsoft Excel. Where multiple papers pertaining to the same study population and site are encountered, we extract data from the most informative paper(s) but they will be linked together as one unique study paper. The earliest year of publication adopted if multiple papers of the same study are published at different times.

3.3.7 Data Items

The data captured include the publication details, language of the paper, study period, study location, geographic setting, study design, study period, characteristics of participants, sample size and sampling technique, explanatory and outcome variables, data analysis and the major findings. Specific data such as the demographic of age and gender also extracted.

The definition of risks, definition and method of measuring outcome used by original author were recorded. For behavioural risk factors of NCDs, the risks were categorised into smoking, alcohol consumption, physical inactivity, and unhealthy diet. The metabolic risk factors of NCDs included increased BP, high blood sugar, abdominal obesity, abnormal triglyceride, or high-density lipoprotein (HDL)-cholesterol levels. The prevalence data on metabolic syndrome was also extracted.

3.3.8 Risk of Bias in Individual Studies

The author and Dr. Nadia binti Mohd Efffendy independently assessed the quality of the included studies before extracting data from them. For quality assessment, the modified Newcastle-Ottawa scale (Wells et al., 2001; Modesti et al., 2016) was adopted to assess the selection, comparability, and outcome of the cross-sectional studies (Table 3.3). These data were presented descriptively in a tabulated form.

The Newcastle-Ottowa Scale was used to assess the study selection, comparability of study groups, and ascertainment of the outcome of selected studies (Wells et al., 2001; Modesti et al., 2016). The maximum score for study quality was ten. This scale is one of the most commonly used worldwide for the assessment of study quality in systematic review and meta-analyses. A previous systematic review by Cook & Reed (2015) concluded the usefulness and reliability of this tool. Nevertheless, this validated instrument still had several limitations, such as the lack of comprehensive manuals and low agreement between two independent reviewers (Luchini et al., 2017). Therefore, in this review, a discussion was conducted with a third reviewer if a consensus agreement needed.

Table 3.3: Newcastle-Ottawa Quality Assessment Scale for cross-sectional studies

Selection(Maximum 5 stars)			
Representativeness of the sample	Sample size	Non-respondents	Ascertainment of the exposure (risk factor)
a. Truly representative of the average in the target population. (all subjects or random sampling) * b. Somewhat representative of the average in the target population. (non-random sampling) * c. Selected group of users. d. No description of the sampling strategy.	a. Justified and satisfactory.* b. Not justified.	a. Comparability between respondents and non-respondents characteristics is established, and the response rate is satisfactory. * b. The response rate is unsatisfactory, or the comparability between respondents and non-respondents is unsatisfactory. c. No description of the response rate or the characteristics of the responders and the non-responders.	a. Validated measurement tool.** b. Non-validated measurement tool, but the tool is available or described.* c. No description of the measurement tool.*
Comparability(Maximum 2 stars)			
The subjects in different outcome groups are comparable, based on the study design or analysis. Confounding factors are controlled.			
a. The study controls for the most important factor (select one).** b. The study control for any additional factor.*			
Outcome (Maximum 3 stars)			
Assessment of the outcome		Statistical test	
a. Independent blind assessment. ** b. Record linkage. ** c. Self report. * d. No description.		a. The statistical test used to analyse the data is clearly described and appropriate, and the measurement of the association is presented, including confidence intervals and the probability level (p value). * b. The statistical test is not appropriate, not described or incomplete.	

Source: Modesti et al. (2016)

3.3.9 Summary Measures and Synthesis of Results

The characteristics of the study participants summarised using narrative synthesis in tables using descriptive statistics. The main outcomes was the prevalence data behavioural and metabolic risks of NCDs. Due to the scarcity of literature on this topic, no studies were excluded based on concerns about sampling bias, representativeness, or sample size. Instead, the heterogeneity in the methodology

quality and the disparity in the results between studies were highlighted. The (weighted) mean prevalence rates found in the review were calculated as follows (Loef, & Walach, 2012; Aryal & Wasti, 2016):

$$\frac{\text{Sum of the number of cases in all studies considered}}{\text{Sum of the number of respondents in all studies considered}} \times 100$$

Finally, the findings were synthesised and summarised in a narrative form in the following finding and discussion section.

3.4 Methods of the Quantitative Phase

To achieve the study objectives of the quantitative research, the five-step process of collecting quantitative data was performed following Creswell (2014). Each step is discussed in detail in the next sections.

3.4.1 Respondents

3.4.1.1 Study Design

The cross-sectional study design of data collection used for this study. Through this method, data can be collected at a specific time. Data can also be measured by multiple variables individually, including exposure to risk factors (Strom, 2012). Also, a cross-sectional design study involves a more representative sample and provides valid information and is also cost-effective (Creswell & Hirose, 2019). Because of the research's objective of providing baseline data on NCDs and its risk factors, as well as their KAP status and health-seeking behaviour, a cross-sectional study at this point was considered appropriate.

3.4.1.2 Study Area

The study took place in the rural district of Jelebu, Negeri Sembilan from January 2018 to April 2019 (Figure 3.3). The map was generated using Google Earth Pro for Windows (Google LLC; United States). Jelebu is the second-largest district in Negeri Sembilan with an estimated total population of 37 287 (DoSM, 2011). The economic activities for the local people are mainly agricultural and small businesses. The Orang Asli from the Temuan group is the main Orang Asli tribe living in Jelebu with estimated 7.5% from total Jelebu populations (JAKOA, 2015).

Figure 3.3: Map of the Jelebu District in Negeri Sembilan, Malaysia

3.4.1.3 Samples

In the sampling process, a sample is a subset of the population that can represent the entire population (Banerjee & Chaudhury, 2010). From the study

population, the geographical location, age, sex, and other attributes and variables such as occupation, religion, and ethnic group can be derived (Banerjee et al., 2007). For the purpose of inferential analysis, administrative boundaries such as districts or states can be suitable to be used as a population. By using geographical location, the study findings can provide useful information for the local administrative authorities and the study respondents. Additionally, the population demographic data such as population size, age, or gender can be easily obtained from the administrative authorities (Banerjee & Chaudhury, 2010).

The demographic characteristics that define the target population, the broad set of people to which the results of the study will be generalised. In this thesis, the target population was adult Orang Asli whereas the study sample was the Orang Asli communities living in Jelebu, Negeri Sembilan. The subtribe ethnic group of Orang Asli living in Jelebu is Temuan from the Proto-Malay group. By choosing a specific ethnic group, it allowed us to compare the study findings with a previously reported study among this particular tribe. Figure 3.4 illustrate the study population and the sample of interest in this study.

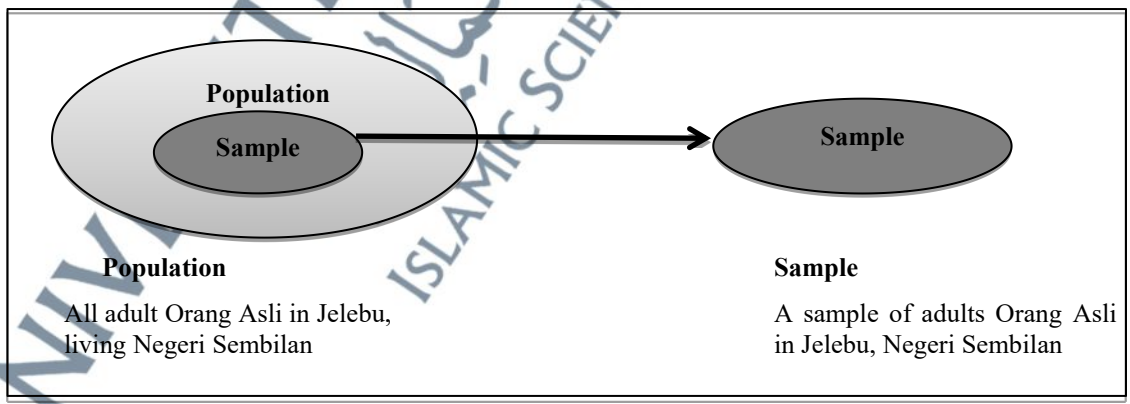


Figure 3.4: Populations and Sample

3.4.1.4 Inclusion and Exclusion Criteria

The inclusion criteria for this study are:

i. Adults aged 18-years-old and above

ii. Malaysian citizens with Orang Asli ethnicity

iii. Able to understand and speak in Bahasa Malaysia

iv. Residents of Jelebu district for at least one years

The samples would be excluded from this study if they:

i. Had psychiatric illness with communication disabilities

ii. Pregnant women

iii. Refused to participate

3.4.1.5 Sample Size Determination

For the calculation of the required sample size, formula provided by Lwanga & Lemeshow (1991) to estimating a populations proportion with specified absolute precision used. The information for the calculation is shown in Table 3.4.

$$n = \frac{Z^2_{1-\alpha/2}\, P\,(1-P)}{d^2}$$

Where,

n = sample size
$z^2_{1-\alpha/2}$ = standardised normal deviation corresponding to the level of significance
P = Prevalence of interest (from literature survey) expressed in percentage form
d = Absolute precision

Table 3.4: Information for the Calculation of the Sample Size using Lwanga & Lemeshow (1991)

Symbols	Value
$z^2_{1-\alpha/2}$	Value 1.96, using 95% CI
P	33.4%. Based on the prevalence of overall hypertension among adult Other Bumiputras ethnicity in Malaysia (IPH, 2015b)
d	0.05

The sample size was then inflated to consider a 10% non-response rate. Thus, from the required sample size as calculated below, the minimum number of respondents required for the study estimated was 376.

$$n = \frac{1.96^2 \text{ x } 0.334\text{x } (1\text{-}0.334)}{0.05^2}$$
$$= 341.82 + (341.82 \text{ x } 0.2)$$
$$= 376$$

The calculation was verified using formula for finite population sample size by Daniel (1999). The information for the calculation is shown in Table 3.5.

$$n = \frac{NZ^2_{1-\alpha/2}\, P\,(1-P)}{d^2\,(N-1) + Z^2 1\text{-}\alpha/2\, P\,(1-P)}$$

Where,

n = sample size
N = Population size
$z^2_{1-\alpha/2}$ = standardised normal deviation corresponding to the level of significance
P = Prevalence of interest (from literature survey) expressed in percentage form
d = Absolute precision

Table 3.5:Information for the Calculation of the Sample Size using Daniel (1999)

Symbols	Value
N	2806. Based on total number of Orang Asli population in Jelebu, Negeri Sembilan (JAKOA, 2010)
$z^2_{1-\alpha/2}$	Value 1.96, using 95% CI
P	33.4%. Based on the prevalence of overall hypertension among adult Other Bumiputras ethnicity in Malaysia (IPH, 2015b)
d	0.05

The sample size was then added to consider a 10% non-response rate. Thus, from the required sample size as calculated below,

$$n = \frac{2806 \text{ x } 1.96^2 \text{ x } 0.334\text{x } (1\text{-}0.334)}{0.05^2\,(2806\text{-}1) + 1.96^2\text{x } 0.334(0.334)}$$
$$= 304.80 + (304.80 \text{ x } 0.1)$$
$$= 335$$

The required calculated sample size using Daniel (1999) was 304.8 and it was added with a 10% non-response rate for a sample size of 335. Therefore, the highest calculated sample size for this study was 376.

3.4.1.6 Sampling Method

A list of Orang Asli villages was obtained from the JAKOA, as shown below in Table 3.6. The Orang Asli villages categorised under the fringe settlement group by JAKOA (2010). They have good accessibility to amenities such as piped water, electricity supply, and connecting roads to their surrounding communities. Additionally, there is also a land development project and a sustainable economic source for them.

Based on the provided list, the study location was selected using simple random sampling. From thirteen Orang Asli villages in Jelebu, eleven villages were selected as shown in Figure 3.5. Following that, the respondents were selected using convenience sampling based on the pre-defined inclusion and exclusion criteria. The researchers selected respondents who were available during the time of data collection and willing to join the study.

Table 3.6: List of Orang Asli Villages in Jelebu, Negeri Sembilan

No	Villages name	State legislative assembly	No. of the household head	No. of the household	Total	Distance to Jelebu Hospital (km)
1	Kampung Orang Asli Tohor	Chennah	89	266	355	53
2	Kampung Orang Asli Bertam	Chennah	23	101	124	40
3	Kampung Orang Asli Jeram Lesung	Chennah	30	63	93	36
4	Kampung Orang Asli Dusun Kubor	Chennah	89	257	346	25
5	Kampung Orang Asli Ulu Kelaka	Klawang	63	130	193	11
6	Kampung Orang Asli Ulu Kemin	Klawang	17	68	85	6

Table 3.6, continued

No	Villages name	State legislative assembly	No. of the household head	No. of the household	Total	Distance to Jelebu Hospital (km)
7	Kampung Orang Asli Baner Tengkoh	Pertang	39	67	106	27
8	Kampung Orang Asli Parit Gong	Pertang	120	383	503	26
9	Kampung Orang Asli Batu 17	Pertang	58	131	189	28
10	Kampung Orang Asli Sawah	Pertang	51	150	201	30
11	Kampung Orang Asli Putra	Pertang	81	189	270	26
12	Kampung Orang Asli Ulu Lakai	Pertang	44	78	122	37
13	Kampung Orang Asli Rangoi	Chennah	52	167	219	45
	TOTAL		**756**	**2050**	**2806**	

Source: Bahagian Pembangunan Modal Insan & Kebajikan (JAKOA Negeri Sembilan & Melaka)

Figure 3.5: Sampling Location of Respondents

According to Creswell (2014), even though convenience sampling is non-probability sampling, the sample can still provide useful input to answer research questions and hypotheses. Non-probability sampling is the commonly used sampling method in recruiting Orang Asli respondents. Due to the limited access to their localities, the shy behaviour of the Orang Asli, and limited number of Orang Asli available during the time of the interviews, it was often hard to reach the people to be selected. Thus, it is highly challenging to achieve the sample size via probability sampling. Therefore, purposive, convenience, snowball sampling, or a combination of these methods are commonly used (Ashari et al., 2016; Mokhsin et al., 2018).

3.4.2 Study Approval

3.4.2.1 Ethics

Approval for the study was obtained from the Research Ethics Committee, USIM (USIM/JKEP/2017/27) and the Malaysia MOH ethical committee from the Medical Research and Ethics Committee (MREC) (KKM/NIHSEC/P18-2338(11), available in Appendix 4 and 5, respectively.

The purpose and procedures of the study were explained to the authorities of JAKOA. The JAKOA granted permission and provided supporting documents to the research team to conduct the study (Appendix 6). The research team also obtained permission from the Jelebu Health District Office.

During an examination, respondents with a raised blood pressure and abnormal level of blood glucose were advised to seek medical attention or treatment at the nearest healthcare facilities.

3.4.2.2 Funding

The research funding provided by the Mizan Research Centre (MRC), Universiti Sains Islam Malaysia. MRC focuses on preserving the Mizan (balance) nature of human beings and its environment. It strives to inculcate the correct understanding of the issue of development and future sustainability (Mizan Research Centre, 2020).

3.4.3 Research Instruments

The instruments used in the Phase Two quantitative study including information sheets, consent forms, and data collection sheets were tested before data collection. The data collection sheets used was a structured questionnaire in the form of booklet.

According to Creswell (2012), the development of study instrument involves four phases, namely planning, development, quantitative evaluation, and validation. All the items used the adopting and adapting method of existing instruments in the development process to build up the questionnaire to answer the research questions. The questions were either originally in Bahasa Malaysia or translated to Bahasa Malaysia. A set of measurement items from the constructs established based on the analysis of the related literature. The item should meet the definition of each construction and have acceptable validity and reliability.

3.4.4 Validation Phase

For the validation of the instruments, the validity and reliability of the tools were determined (Creswell, 2012). Validity is defined as the extent to which a

concept is accurately measured in quantitative research whereas reliability is the accuracy of an instrument (Heale & Twycross, 2015).

3.4.4.1 Content Validity

The content validity looks at whether the instrument adequately covers all the necessary content concerning the variable. A subset of content validity is face validity in which experts are asked about their view about whether an instrument measures the construct intended (Heale & Twycross, 2015). The initial draft of the questionnaire was given to different experts from various background of professions for review, including:

i. Three public health specialists
ii. Two nutritionist
iii. Two medical officer from Jelebu District Health Office
iv. One staff nurse from Klinik Kesihatan Jelebu
v. Two health attendant from Jelebu District Health Office of Orang Asli ethnicity

After reading each of the item construct, the experts commented on each item and gave their opinion whether it measured the intended concept. After that, the questionnaire was reviewed and edited based on the comments and suggestion from the experts. Table 3.7 shows the feedback given by the experts. Several items were reworded for to clarify their meanings better to the laypeople.

Table 3.7: Comments and Decision from Expert Feedback

No	Comments/Recommendation	Decision
1.	Diet change to *pemakanan*	Replaced
2.	Kentang change to other carbohydrate e.g noddles	Replaced
3.	Add *lumpuh separuh badan* next to *angin ahmar* in the knowledge section	Added
4.	Bold each NCDs category in the knowledge section	Accepted and modified
5.	Bold each lifestyle behaviour in the practice section	Accepted and modified
6.	The font is small	Font 10 changed to 12
7.	Add check list at the beginning of the questionnaire	Added

3.4.4.2 Construct Reliability

A pilot test was conducted at Kampung Orang Asli Kemensah, Selangor. A total of 36 responses were collected from the Orang Asli of the Temuan tribe in this village. They were chosen as a the samples in the pilot study as they had a similar background with the study population. All of the respondents from the pilot test were not included in the main study.

In general, their mean age was 39.90 (*SD* = 14.63). The majority of them were female (70.0%), married (83.3%), with primary school education level (50.0%). There were no outliers, and the data were suitable for further analysis. The mean time for each interview session was 22.30 minutes (*SD* = 8.76).

Homogeneity of internal consistency was assessed using Cronbach's alpha (Heale & Twycross, 2015). The coefficient of Cronbach's alpha reliability ranges between 0 and 1. The closer it is to 1.0, the more significant the internal consistency of the items would be on the scale (Gliem & Gliem, 2003). George & Mallery (2003) provided the following rules of thumb for Cronbach's alpha value, i.e. 0.90 or higher is considered as excellent and 0.50 or less is considered as unacceptable.

Table 3.8 shows the Cronbach's alpha value for each construct. In short, all the Cronbach's alpha values were acceptable. Since further inspection of the item-to-total correlation for all the scales showed higher values than the recommended score

of 0.50 (George & Mallery, 2003), all the items were included in the final questionnaire.

Table 3.8: Cronbach Alpha Value

Construct	Number of items	Cronbach Alpha Value
Knowledge		
i. General and example of NCDs	7	0.847
ii. Heart attack	4	0.938
iii. Stroke	4	0.955
iv. Hypertension	4	0.905
v. DM	4	0.842
vi. COPD	4	0.993
vii. All items	27	0.961
Attitude		
i. NCDs behavioural risk	6	0.612
ii. NCDs management	6	0.860
iii. Support network	3	0.885
iv. All items	15	0.898
Practice	5	0.705

3.4.5 Study Independent Variables and Measurement

3.4.5.1 Sociodemographic and Socioeconomic

a) **Age:** The respondents' current age calculated at the time of the data collection date. Adults aged 18 years old and above during the last birthday as stated in Malaysian identification card. Based on Bujang et al. (2012) and Ahmad, Nawi & Madihah (2015), the adult group is categorised into:

   i. Young adult : aged 18 to 39 years old.

   ii. Elder adult : aged more than 40 years old.

b) **Gender:** Male or female as stated in the Malaysian identification card.

c) **Ethnic group.** Race as stated in the Malaysian birth certificate. For this study, only Orang Asli Semenanjung were included.

d) **Period of stay.** The respondents' period of stay in Jelebu. Only respondents staying longer than one years were included as they would be familiar with the healthcare facilities provided by the government.

e) **Marital status.** The lawful recognition of the relationship or agreement between a man and a woman, to be husband and wife. They was categorised (IPH, 2015a) as follows:

i. Single : those who have never been married.

ii. Married : those who are still married at the time of the interview.

iii. Widowed : those who are not remarried after the death of the spouse during the interview.

iv. Divorced/separated permanently: those whose marriage has been annulled through a divorce or religious law or who have not lived together for a long period and are unlikely to reunite.

f) **Formal education.** Any formal highest Malaysian education received. The education is categorised (IPH, 2015a) into:

i. No formal education : those who have never attended any formal educational institution.

ii. Primary education : those who have attained the highest level of education in years 1 to 6 or equivalent.

iii. Secondary education : those who have achieved the highest level of education in Form 1 through Form 5.

iv. Tertiary education : those who have attained the highest education after Form 5.

v. Unclassified education : Others.

Those with no formal education or primary education were categorised into low education group whereas those with secondary education or above were categorised into high education group (Wong et al., 2018).

g) **Working status.** People who work at least one hour during the referral week to earn wages, benefits, or family benefits whether as employers, employees, self-employed, or non-paid family workers (DoSM, 2019).

h) **Employment status.** Classified into two categories of working and not working. For those who were working, they were further categorised into government employee, government-linked companies (GLC) employee, non-government employee, and self-employed. For those who were not working, they were categorised into student, housewife, retirees, unemployed, and others.

i) **Monthly household income.** Overall income earned by household members, whether in cash or kind and can be referred to as gross income on Malaysian Ringgit currency. The respondent was categorised based on their Poverty Line Income by Economic Planning Unit, Prime Minister's Department in 2016 (Ministry of Housing and Local Government, 2020). A Peninsular Malaysian whose household monthly income is below RM607 a month is categorised under the hardcore poverty group.

3.4.5.2 NCDs Behavioural Risk Factors

The respondents were asked on NCDs behavioural risks of smoking, alcohol consumption and physically inactive. A code of 1 was assigned if the respondent answered "yes" and the code of 0 was assigned if the respondent answered "no".

a) **Smoking.** Current smokers defined as respondents who smokes any tobacco product daily or occasionally during the study (IPH, 2015b).

b) **Alcohol consumption.** Intake of any drink containing ethanol irrespective of concentration and inclusive of those consumed for medical purposes such as alcoholic tonic for the past twelve (12) months before the interview (IPH, 2015b).

c) **Physically inactive.** The CDC (2020) definitions for exercise is a type of physical activity that involves planned, structured, and repetitive bodily movement has done to maintain or improve one or more components of physical fitness. In this study, physical activity refers to activity undertaken for fitness, recreation or sport that caused a large increase in the respondent's heart rate or breathing (Milton, Bull & Bauman, 2011; Australian Bureau of Statistics,2014).

This variable was derived from the quantitative survey question: "In a typical week, did you spend more than 30 minutes for at least three times per week doing physical activities outside of work that result in an increase in your heart rate and breathing? This may include sport, exercise, and brisk walking for recreation or to get to and from places, but should not include housework or physical activity that may be part of your job". The respondent was considered physically inactive if they reported less than a minimum of 30 minutes at least three days per week of moderate-to intense intensity physical activity (Myers, 2003).

There are no reliability or validity data available on this single item question from the physical activity; however, previous research has shown acceptable levels of validity and reliability of recall of moderate- to vigorous-intensity physical activities including in study among indigenous peoples (Milton, Bull & Bauman, 2011; Doggett & Dogra, 2015).

d) **Inadequate vegetables intake.** The Malaysian Dietary Guidelines defined adequate vegetables intake as consuming at least 3 servings of vegetable per day (NCCFN, 2010).

3.4.5.3 Obesity Indicators

a) **Abdominal obesity.** A measuring tape was used to measure the waist of the subject. The measurement was done at a level midway between the lower rib margin and the highest point of the iliac crest with the tape placed all around the body in the horizontal position. The measurement was done with the subject standing upright in a stable position. The measurement was taken at the end of a normal expiration (WHO, 2000). Each measure was repeated twice and the nearest one-centimetre reading was recorded.

The cut-off point for abdominal obesity was based on the recommendations by the WHO Expert Consultation on Obesity (2011) and the IPH (2015b). In this classification, the waist circumference of more than 90 cm for male and more than 80 cm for female were considered as abdominally obese.

b) **Body mass index (BMI).** BMI was calculated using the respondent's height and weight. Height was taken with no shoes to the nearest centimetre using a digital SECA 703 column scale (seca GmbH & Co. KG; Germany) with a front-mounted height rod. Wearing only light clothing and no shoes or items in their pockets, the respondents stood on a digital SECA 703 column scale that was placed on a hard horizontal tiled floor to measure their weights. Two measurements were taken in succession to ensure accuracy and the result was recorded to the nearest half kilogram. Then BMI was calculated using the following formula:

$$\text{BMI} = \frac{\text{Weight (Kg)}}{\text{Height squared (m}^2\text{)}}$$

Obesity for the Asian population was used to define BMI obesity using the WHO Asia Pacific guidelines (WHO, 2000; Weisell, 2002). BMI was classified according to the WHO BMI cut-off for Asian as follow:

i. <18.5 kg/m^2 : Underweight

ii. 18.5-22.9 kg/m^2 : Normal

iii. 23.0-24.9 kg/m^2 : Overweight

iv. ≥ 25.0 kg/m^2 : Obesity

c) **Body Fat Percentages (BFP).** Body composition was assessed using bioelectrical impedance analysis (BIA) performed with Maltron BF-906 Body Fat Analyser (Maltron International Ltd; Essex UK) (Appendix 7). Check-list for standardized BIA procedures by González-Correa & Caicedo-Eraso (2012) was used. Before the measurement, hydration, physical activity status and fasting status were assessed. Then, the respondents were asked to remove any jewellery and item made of metal including metallic watch, belt, necklace and ring.

Two electrodes were placed on the hands and another two on the feet. After entering all the required parameters and having the subject at rest in a supine position with hands and legs slightly apart, the Maltron BF-906 was activated.

This tool measures the flow of electrical signals as they pass through fat and lean areas as well as water in the body. When the amount of fat and lean matter or water change, the signals also change to give a reliable and accurate measure of the amount of each of these components that make up the total weight of the person (Salleh et al., 2007).

3.4.5.4 Blood Pressure Measurement

The auto-inflate digital electronic Omron HEM-7120 Automatic BP Monitor (Omron Corporation; Japan) was used to measure the blood pressure. A standard adult-size cuff 145 (wide) x 466 (length) millimetre delivered a cuff pressure of 0-300 millimetre of mercury (mmHg) (+3 mmHg). The automatic blood pressure monitor was used due to its ease of transportation and safety. A study by Myers et al. (2011) concluded that automatic blood pressure could be used in primary care practice to obtain accurate readings.

Blood pressure measured were taken when the subject is calm, not sought after meals or after exercise and after at least 15 minutes of rest. The reading taken during sitting on the left hand supported at heart level in sitting upright position. It was measured twice, with two minutes apart when the respondent is at rest. Then, the average measurements used as the BP value for individual respondent (Imai et al., 2003). In this study, high blood pressure was considered as greater than 140mmHg systolic and/or 90 diastolic mmHg (IPH, 2015a).

3.4.5.5 Blood Glucose Measurements

Blood glucose was measured using glucometer from Accu-Chek Performa meter (ROCHE Holding AG; Basel, Switzerland). The glucometer was first calibrated by using the chip code of the respective test strips. Firstly, the code key was inserted into the test strip slot. Then, the number appearing on the screen was checked that it was the same as that on the box of test strips. The finger site was first cleaned with an alcohol swab and let dry before being pricked with a lancet. The drop of blood was then put on the test strip before it was inserted into the glucometer. The value of

fasting blood sugar (FBS) of 6.1 mmol/L or more (or random blood glucose of more than 11.1 mmol/L) was considered as abnormal (IPH, 2015a).

3.4.5.6 Health-seeking Behaviour

Health-seeking behaviour is defined as actions taken by individuals with illness in seeking treatment by self-medication, purchase of medicines from pharmacies, seeking treatment from traditional healers, government, and/or private healthcare facilities (IPH, 2015b). A series of the question was asked to the respondents on their action in seeking treatment and reason for not seeking treatment if they did not do so. In addition, any information on any traditional medicine used to heal any illness was also collected. The questions (Table 3.9) was adopted and adapted from methodology used in NHMS 2015 (IPH, 2015a).

Table 3.9: Health-seeking Behaviour Items

Components	Items/ questions	Respond/ scoring
a) General health-seeking behaviour	If you are sick, do you go for treatment?	Yes/ No
b) Reason for not seeking treatment	If not, what is the reason?	List of options given (one only) i. Unable to bear the high treatment costs ii. Do not believe / like the services provided by health facilities iii. Distance/transportation problem iv. Don't know where to go v. Not enough time / busy with work vi. Never bother about the pain
c) Selection of treatment	If you seeking for treatment, what is your treatment of choice	List of options given i. Modern ii. Traditional iii. Both
d) Selection of modern healthcare facility	State the most common modern healthcare facility to visit for treatment.	List of options given i. Government clinic ii. Government hospital iii. Private clinic iv. Private hospital v. Pharmacy
e) Selection of traditional treatment	State the traditional treatments that you often use to overcome your health problems.	List of options given i. Traditional massage ii. Islamic medical practitioners iii. Traditional healer/ *Tok halaq* iv. Other

Table 3.9, continued

Components	Items/ questions	Respond/ scoring
f) Selection of traditional medicine	State the most common traditional medicine used to solve health problems.	List of options given i. Massage or hot oil ii. *Air penawar* or treatment given by Traditional healer/ *Tok halaq* iii. Natural resources iv. Herbs v. Other

3.4.5.7 Knowledge on NCDs

Knowledge in this thesis refers to the assessment of the extent to which an individual understands the concepts regarding major NCDs. It included local knowledge and beliefs, awareness of disease management, awareness of diseases, and the symptoms of diseases. Knowledge about NCDs and risk factors for NCDs was assessed with an adapted, validated and reliable knowledge questionnaire.

In Table 3.10, the 27 individual items, for knowledge sections NCDs and the risk factors, were presented. This question was adopted and adapted from Frantz (2008), Demaio et al. (2011), Ohar et al. (2010), Jones (2017) and WHO (2017). The answers option to knowledge questions were 'Yes' 'No' and 'Don't know.

Table 3.10: Questionnaires for Knowledge on NCDs

Components	Items/ questions	Respond
a. General knowledge	Have your ever heard about Non-communicable diseases?	Yes/ No
b. Examples of NCDs	Which of the following diseases is Non-communicable diseases? i. Diabetes ii. Tuberculosis iii. Asthma iv. Dengue fever v. Heart attack vi. Leptospirosis	Yes/ No/ Don't know
c. Heart attack	Which of the following is a risk factor for heart attack? i. Regular consumption of foods that contain high fat content ii. Balanced diet iii. Smoking iv. Exercise on a regular basis	Yes/ No/ Don't know

Table 3.10, continued

Components		Items/ questions	Respond
d.	Stroke	Which of the following is a risk factor for stroke? i. High blood pressure ii. Exercise on a regular basis iii. Balanced diet iv. Obesity	Yes/ No/ Don't know
e.	Hypertension	Which of the following is true about hypertension? i. A person suffers from high blood pressure if the systolic reading exceeds 140mmHg ii. A person with high blood pressure does not need to take medication as prescribed iii. A person who is obese has a high risk of getting high blood pressure iv. A person with hypertension does not need to undergo regular follow-up treatment	Yes/ No/ Don't know
f.	Diabetes mellitus	Which of the following is true about DM? i. Excessive starchy foods intake such as rice, noodle and bread cause diabetes ii. Someone with family history of diabetes is at risk for developing diabetes iii. Diabetics do not need to take insulin injections iv. Diabetes could not cause kidney damage	Yes/ No/ Don't know
g.	Chronic obstructive pulmonary diseases	Sign and symptoms of COPD? i. Breathless ii. Produce phlegm in large quantities iii. Whistles and sounds breath iv. Chronic cough for more than 3 month	Yes/ No/ Don't know

3.4.5.8 Attitude towards NCDs

The attitude section consisted of fifteen questions that were used to measure the degree of individuals perceived and weighted their behaviours either positively or negatively. The statements was adopted and adapted from Malaysia Clinical Practice Guidelines (CPGs) on management of cardiovascular diseases (MOH, 2017a), stroke (MOH, 2012a), hypertension (MOH, 2018), dyslipidemia (MOH, 2017b), DM (MOH, 2015), obesity (MOH, 2004) and COPD (MOH, 2009).

It consisted of six statements on respondents' perceptions on the NCDs behavioural risks, six statements on disease management, and three statements on their attitude in providing financial and other assistance to family members with NCDs (Table 3.11). The questions were presented as statements and the respondents

were asked to indicate their agreement or disagreement to the statements. In this section, the 5-point Likert scale method was used, ranging from strongly disagree (1), disagree (2), neutral (3), agree (4), to strongly agree (5). The 5-point Likert scale is the preferable scaling item as it is simple, easy to understand, and more suitable for health research (Garratt, Helgeland & Gulbrandsen, 2011).

Table 3.11: Statements for Attitude towards NCDs

Components		Attitude Statement	
a)	Attitude towards behavioural risk factors of NCDs	i.	You feel it is important to practice balanced eating continuously
		ii.	At your age, periodic physical activity is essential to maintaining your health
		iii.	You need to take food following the calories that you need for your daily activities
		iv.	In the selection of foods, you will make sure its low in fat
		v.	If your family members smoke, you will try to advise them to quit smoking
		vi.	You feel it's important to look after your health
b)	Attitude towards NCDs management	i.	If you diagnosed with NCDs, you would try to control the disease before it becomes worse
		ii.	You bring along your medicines when visiting relatives who are far away
		iii.	When you have diabetes, you will not consume foods that are high in sugar
		iv.	Taking the medication/injection provided by your doctor regularly following the prescribed schedule can control diabetes complications
		v.	A person with high blood pressure should undergo a blood pressure check during follow-up treatment
		vi.	If you have high blood pressure, you will reduce your salt intake
c)	Attitude towards support network	i.	You will provide financial assistance if your family member has a stroke
		ii.	You can take care of your family members if they have a stroke
		iii.	You will encourage your female family members to undergo breast cancer screening at the health centre

3.4.5.9 Practice towards NCDs

For the practice of lifestyle behaviour, the measurement refers to the frequency of each lifestyle behaviour. The statements in Table 3.12 was adopted and adapted from National Coordinating Committee on Food and Nutrition (2010) and the WHO (2017). Five questions were used to assess the modifiable behaviours that represented NCDs risk factor i.e. physical activity, vegetable consumption, smoking

behaviour, alcohol intake, and body weighing. The respondents were asked about how frequently they performed each component of lifestyle behaviour using a three-point scale of never (0), seldom (1), and always (2).

Table 3.12: Questionnaires on Practices towards NCDs

Components	Attitude Statement
a) Physical activity	I spend at least 30 minutes exercising 3 or more times a week
b) Smoking behaviour	I am not a smoker
c) Vegetable intake	I eat at least 3 serving of vegetables a day
d) Alcohol behaviour	I never consume alcohol
e) Body weighing	I weigh myself every week

3.4.5.10 Accessibly to Public Healthcare Facility, Services and Medicine

Wilkinson (2001) defined access as the fundamental concepts to the understanding of health-seeking behaviour. Furthermore, according to Grundy & Annear (2010), the action of an individual in seeking treatment is determined by households, health institutions, and their socio-cultural environment. This sections aimed to understand the utilisation, demand for service, perceptions of quality of care, and health belief of the Orang Asli.

The questions in Table 3.13 was adopted and adapted from methodology used in NHMS 2015 (IPH, 2015a). This section consisted of two sub-section. The first section was made up of four items about accessibility to the government healthcare facilities were asked, including a) distance, b) transportation, c) operational hour, and d) waiting time to get treatment. In the second section, another five items were asked about accessibility to get medicine and treatment from government healthcare facilities. The answer options were ‘Yes’, ‘No’, or ‘Unsure’.

Table 3.13: Accessibly to Public Healthcare Facility, Services and Medicine Items

Components		Items/ questions		Respond
a)	Accessibility to government health care centres (distance and time)	i.	The distance between home and the government health care centre is close	Yes/ No/ Unsure
		ii.	Did not experience any problems in terms of transportation to the government health care centre	
		iii.	Operating hours for government health care centres are suitable	
		iv.	Waiting time at the government health care centre is not burdensome	
b)	Accessibility to get treatment and medicines at the government health care centre	i.	Government health care facilities in your area are in good condition	Yes/ No/ Unsure
		ii.	Medicines can be obtained from government health facilities (hospitals / clinics)	
		iii.	Modern medicine is more readily available than traditional medicine	
		iv.	Feel comfortable and will continue to seek treatment at the government health care centre	
		v.	Recommend family members and friends for treatment at a government health care centre	

3.4.6 Study Dependent Variables

For NCDs status, the presence or absence of chronic diseases was based on reported information on chronic conditions diagnosed by a doctor or healthcare professional. During the survey, respondents were asked: "Have you ever been told by a doctor or assistant medical officer (AMO) that you have _____?" followed by a list of chronic diseases including hypertension, hypercholesterolemia, Diabetes mellitus (DM) and others to which the answer could be "yes" or "no". Then, the response obtained from the respondent was confirmed with the records from Jelebu Health District Office. A code of 1 was assigned if the respondent answered "yes" to having any NCDs, while the code of 0 was assigned if the respondent answered "no".

a) **Hypertension.** Reported as being told to have hypertension by a doctor or AMO and evidenced by the patient's medical report (IPH, 2015b).

b) **Hypercholesterolemia.** Reported as being told to have hypercholesterolemia by a doctor or AMO and evidenced by the patient's medical report (IPH, 2015b).

c) **Diabetes mellitus.** Reported as being told to have DM by a doctor or AMO and evidenced by the patient's medical report (IPH, 2015b).

3.4.7 Administering Data collection

Five research assistants recruited for data collection procedure with a minimum tertiary education background. A one day's training and pretest were conducted to exposed the research assistants on to the methods of sampling, interpersonal communication, obtaining informed consent and a survey questionnaire. Training also focused on interactive sessions to introduce data collection methods and correct measuring methods for the anthropometric measure, blood pressure and blood glucose measurement.

All research assistant practices and trained according to the standardized way of WHO guideline. Questionnaires clarified further after the pretest. Guidelines for research assistant provided in Appendix 8. By conducting this training, it allows mitigating the risk of error and bias (Creswell, 2014). Additional to that, the review of the questionnaire carried out at the end of each data collection to minimize error in data collection.

Questionnaire survey was used for the quantitative data collection from the respondents (Appendix 9). By using a standardized set of questionnaires, its minimize variation in data collection. The study purpose, issues of confidentiality, anonymity, and respondents' rights to withdraw from the study at any stage without notice were explained to all the respondents using the respondents' information sheet (Appendix 10). If they agreed, they would sign the informed consent form (Appendix 11).

The study was conducted in compliance with the protocol and standard operating procedure (SOP) of the Clinical Research Centre (CRC). The protocol was

designed to ensure adherence to the ethical principles listed in the Declaration of Helsinki (World Medical Association, 2001). After obtaining the consent from the respondents, face-to-face interview was conducted in the community hall of the selected villages. All the paper-based and electronic copies of the questionnaires were stored in the safe place by the principal investigator.

3.4.8 Data Validation Process

A few precautions were taken to ensure the overall consistency of the data entered, such as establishing the data standard, i.e. codebook, data screening, data cleaning, and missing data management.

3.4.8.1 Data Standard Development

A codebook was built as a standard for the data. This codebook includes sections in the questionnaire, names of variables, variable types, unit and definition of values of the variables. It was used in the database as a reference for coding the responses and missing information. The coding used for the face-to-face questionnaire shown in Appendix 9.

3.4.8.2 Data Screening

To ensure the completeness of the data collected, the interviewers were reminded to complete the data according to the response of the respondents. All questionnaires were reviewed for completeness before the end of the interview. Besides, prior to data entry, each questionnaire was double-checked by the researcher to ensure that no pages and details were missing. Each questionnaire was labelled with a unique identifier number to protect the confidentiality of the identity of the

respondent. Data were entered into the statistical software, Statistical Package of Social Sciences (SPSS) version 23 (IBM Corp; Armonk, New York).

After completing the first data entry, the data enters for the second time as a data checking methods (Barchard & Pace, 2008). In this technique of double-entry data checking, the computer compares the first and second entries to ensure that they match, and also detects values beyond the range permitted (Barchard & Verenikina, 2013). For example, if the data sheets show a three-point agreement scale, entries should be only 0, 1, or 2. The entered data were found to be consistent and accurate with the questionnaire.

3.4.8.3 Data Cleaning

Following the data screening process, data cleaning was performed to identify; (1) duplicate respondents, (2) repeated entry, (3) wrong codes, (4) outliers, (5) missing values, and (6) out of range or extreme observations. Data were analysed using frequency and descriptive statistics. This cleaning process was carried out using SPSS version 23. A copy of the dataset was stored on Google Drive, e-mail, external hard disk and Universal Serial Bus (USB) flash drives for security purposes and to save from any unexpected damage. All information handled with high confidentiality.

3.4.9 Quantitative Data Analysis

3.4.9.1 Descriptive Analysis

The Statistical Package of Social Sciences (SPSS) Version 23 (IBM Corp; Armonk, New York) statistical computer software program was used to analyse all the statistical data. The mean values and standard deviation (*SD*) were included in the descriptive analysis for the continuous variables. The prevalence of major NCDs,

behavioural risks of NCDs, obesity indicators and metabolic risks of NCDs among respondents was calculated in percentages. The data were expressed as numbers and percentages (%) for the categorical variables.

In the knowledge section, correct answers were given one point. For any wrong or unsure response, zero point was given. The total score of knowledge ranged from zero to 28 marks. In the attitude section, the answer option was on a 5-point Likert scale whereby the most strongly positive attitude was given a score of five and the most negative attitude was given a score of one. With 15 statements, the minimum and maximum total scores for attitude were 15 and 75 respectively. Finally, in the practice section, those who answered "always" to the practice were given two points, "seldom" were given one point, and zero for "never". The total score of practices ranged from zero to ten marks.

For categorizing KAP towards NCDs, there was somewhat arbitrary, like many other scales (Conner & Norman, 2005). In this study, the cut points were based on scale, and similar ones used with the KAP towards NCDs module have been previously published and allow for the standardized comparison of knowledge, attitude and practices scores (Ithnin et al., 2018; Ithnin et al., 2020). To determine the level of knowledge, attitude, and practice, the total score of each individual respondent was converted into percentages. Any scores of less than 74% indicated poor KAP categories, whereas scores of more than 75% indicated good KAP categories (Semungus, Tafese & Semella, 2017; Chong, Appannah & Sulaiman, 2019; Van Hooser et al., 2020).

3.4.9.2 Inferential Analysis

Inferential analysis for hypotheses testing made using the Pearson correlation test, and univariate and multivariate binary logistic regression test.

Pearson correlation analysis were performed to check the relationship between knowledge score with attitude and practice scores, and also between attitude and practice. According to Chan (2003), correlation coefficient value, *r* of at least 0.8, indicating very strong relationship, 0.6 to 0.8 consider moderately strong. Whereas, *r*-value of 0.3 to 0.5 and less than 0.3 showing the fair and poor relationship, respectively. The *r*-value of 0 indicating none relationship. The results were significant if *p*-value < .05.

A logistic regression test was conducted to identify which independent variables could predict an outcome of dependent variables. The binary logistic regression was used as an outcome variable since the data is a categorical dichotomous outcome.All dependent variables are binary, where "1" represents outcome and "0" for the reference group.

The univariate analysis was first carried out to compare each variable to the outcome individually. Chi-square value from likelihood ratio, crude odds ratios, 95% confidence interval (CI), and p-values were reported in the tables. Following that, a linear regression test was performed to determine the multicollinearity between the dependants variables of interest and the associated factors. There is no multicollinearity if the standard variance inflation factor (VIF) values were less than ten, tolerance values exceeded 0.1, and the Durbin-Watson value between the two critical values was 1.5 to 2.5 for independence assumption (Chan, 2004a).

Multivariate binary logistic regression test was then conducted to find out the most significant predictors that influenced the dependent variables. The binary

logistic regression was used as an outcome variable since the outcome data was a categorical dichotomous variable. All the variables whose *p*-value < .25 along with the variables of known importance contributor include in the analysis (Agresti, 2007). Model validation was done using the standard Hosmer-Lemeshow test for goodness of fit (Chan, 2004b).

The logistic coefficient (*B*), Wald's statistic, exponentiation of the B coefficient (*Exp (B)*) for adjusted Odds Ratios (*aOR*), 95% CI, and *p*-values were reported in tables. The *B*-value is the logistic coefficient for each predictor variable for each alternative category of the outcome variable, not the reference category. The *B*-value is the expected amount of change in the logit for every one-unit change in the predictor. The closer the value of *B* is to zero, the less influence the predictor has in predicting the logit. The Wald value estimates give the "importance" of the contribution of each variable in the model. The higher the value, the more "important" it is. The *Exp (B)* is the adjusted Odds Ratio (*aOR*) associated with each predictor. Predictors that increase the logit will display *Exp (B)* greater than 1.0 whereas predictors that do not affect the logit will show an *Exp (B)* of 1.0 while the predictors that decease the logit will have *Exp (B)* values of less than 1.0 (Chan, 2004b). The analyses were set at a significance level of $p < .05$.

3.5 Methods of the Qualitative Phase

The respondents for the qualitative phase were selected among the total number of respondents (*n*=325) who completed the quantitative phase of the study and had consented to be included in the follow-up interviews. The best respondents to interview are those who will deliver the richest information (Patton, 2002). By

purposefully selecting the respondents for the semi-structured interviews in this study, a deeper understanding and comprehension could be obtained regarding their situation.

3.5.1 Study Design

The qualitative in-depth interview was used to study individual transition, health behaviour, knowledge and barriers, also perception and thoughts for accessing or seeking treatment behaviours. This approach helps to explain the reasons for the outcome from the quantitative research.

3.5.2 Setting

From the thirteen Orang Asli villages in Jelebu, Negeri Sembilan, one Orang Asli village was selected as the study area of the qualitative study. The village were purposely selected as they fulfilled the criteria that enabled the researcher to understand the critical issues related to NCDs. The communities have also been chosen to examine the possible effects of urbanisation on NCDs, proximity to health services, and the usage of modern healthcare services.

The selected communities were located in Kampung Orang Asli Ulu Kelaka, Kuala Klawang (Coordinates: 2°53'41"N, 102°2'27"E). There was good premixed road connecting the communities to other villages and towns. This village was a new settlement built in 2003. Figure 3.6 and 3.7 shows the location of Kampung Orang Asli Ulu Kelaka and the nearest government clinic and hospital, namely Klinik Desa Ulu Klawang (1.7 km) and Hospital Jelebu (11.2 km).

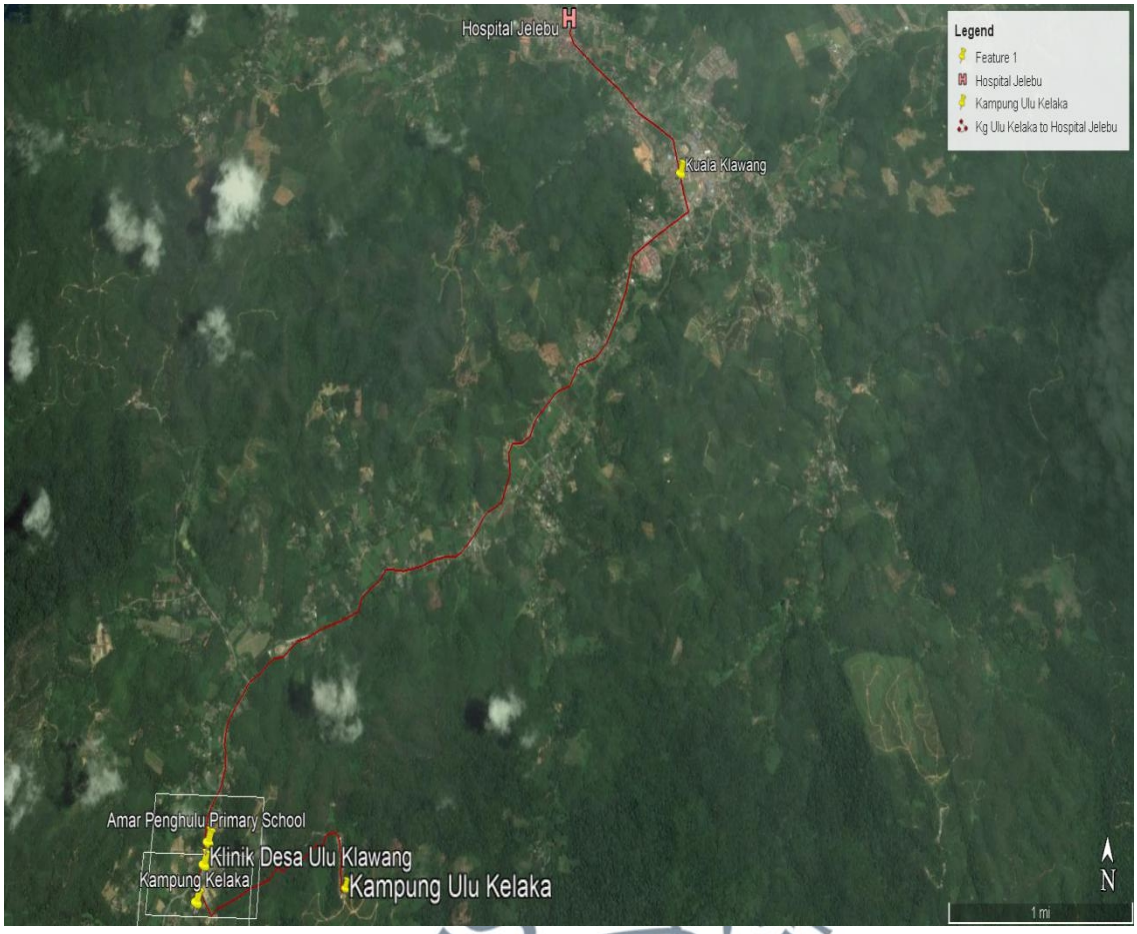


Figure 3.6: Kampung Orang Asli Ulu Kelaka, Klinik Desa Ulu Klawang and Hospital Jelebu location

Figure 3.7: Kampung Orang Asli Ulu Kelaka from Google Earth

3.5.3 Ethical Consideration and Approval

The ethical consideration was included as part of the quantitative phase during the ethical clearance process. The ethics committees from the MREC, Ministry of Health Malaysia (Appendix 5) and JAKOA (Appendix 6) were aware that the study has quantitative and qualitative phases. Before the qualitative interview, the Tok Batin (Chief of Orang Asli) and the chairman of the Village Community Management Council (MPKK) of Kampung Orang Asli Ulu Kelaka were informed. They granted the permission to conduct the study.

The potential respondents were contacted by telephone or via the representative of the village to provide them with information regarding the qualitative research. If they declined to take part, the next person on the list would be invited. If they agreed to take part, an appointment would be given to the respondents regarding the time and the venue for the interviews. This step was repeated until all the interviews were completed.

No money or compensation was offered to respondents during recruitment. However, they were given a token of appreciation that included a pack of rice, noodles, and ketchup for their daily consumption. All those who participated in Phase Three were thanked for volunteering and informed that their participations were highly valuable towards this study.

3.5.4 Recruitment of Respondents for the Interview

Purposive sampling was performed in which the best respondents that could address the research question were chosen (Rowley, 2012). During the quantitative phase, respondents were informed about the probability of a second data collection which was the qualitative process. To ensure the maximum variation of respondents,

different genders, education, household income range from low to high incomes and health status have been selected.

The chairperson of Komuniti Sihat Perkasa Negara (KOSPEN) shall be the contacted representative of the village. The researcher contacted the contacted representative and tell about the respondents who met the interview requirements. To get their verbal consent for the in-depth interview, the village representative approached the respondents via a phone call or meeting in person. The purpose and confidentiality of the interview was explained to the respondents. Only those who consented verbally were invited to the interviews. The date was set according to the availability and preference of the participant. The contacted representative was only informed of the number of interviewees eligible and the date of the interview. However, the information collected on the interviews was kept confidential.

3.5.5 Sample Size

The sample size of qualitative research is often smaller than that used in quantitative research. This is because in-depth interview for qualitative research is more inductive and emergent in its process. Unlike quantitative research, it does not tend to rely on hypothesis testing as there is less concern to make generalisation to a larger population of interest (Dworkin, 2012).

According to Trotter (2012), the ideal standard for the sample size in qualitative research is to saturation in the interview process. The saturation is a point at which all questions have been thoroughly explored in detail and no new concepts or themes emerge in subsequent interviews (Saunders et al., 2018). Therefore, each qualitative interview transcript should be transcribed after each interview session. Once the data is saturated, the interview process can be terminated.

3.5.6 Research Instruments

The interview process consisted of three main sessions, namely introduction, actual interview, and closing (Appendix 12). The topic guide was developed base on the preliminary findings of the quantitative phase.

In the introduction session, the respondent was informed about the study details and they were given assurance about the ethical principles applied in the study, such as anonymity and confidentiality. They were also told that the interview would be audio recorded. This session provided the respondents with some idea of what to expect from the interview. It could also improve the trust of the respondents, thus serving as a fundamental aspect of the informed consent process. They were also given a study information sheet (Appendix 13). The interviewer would only proceed after obtaining the respondents agreed to participate in the survey and signed the consented form (Appendix 14).

All this information was gathered during the actual interview. The topic and the directions of questions employed to achieve the qualitative findings are presented in Table 3.13. The main questions were semi-structured open-ended and probing. The topics of the transitional change, knowledge on NCDs, the risk behaviour towards NCDs and healthcare-seeking behaviour explored in more detail.

At the end of the interview, the interviewer thanked the respondents for their time. They also asked if the respondents have anything they would like to add. By doing so, the respondents could outline any issues that they have thought about or felt were important but have not been mentioned by the interviewer (Gill et al., 2008). The interviews took about 30 to 50 minutes.

The saturation point was monitored continuously throughout the recruitment. After the 14th interview, there was no new theme emerging from the interviews.

Therefore, it was considered that the data collection had reached a saturation point. However, data collection was continued for two more interviews to confirm that no further new themes.

Table 3.14: The Topic and Directions of Questions of the Qualitative Study

Topic		Directions of Questions	
a)	Resettlement to the village, development and life transition	i. ii. iii.	Their past history as in whether they moved from elsewhere, where were they born, where did they grow up. How did the transition affect them? Has the change of environment altered their lifestyle / diet / activities / health status?
b)	Knowledge and perception on NCDs and health effects of NCDs risk factors	i. ii.	To underpin the knowledge on NCDs and its risk behaviours To assess their understanding on negative effects and risks of unhealthy behaviour on health
c)	Barriers to preventing NCDs risk factors	i. ii. iii. iv.	To assess the influencing factors smoking behaviour and barrier to stop smoking. To assess the the influencing factors of alcohol consumption and barrier to stop alcohol intake behaviour. To assess any diet change and the factors that influence their consumption of fruits and vegetables in terms of social, economic, geographical and accessibility To assess the experience in terms of challenges or difficulties and support in carrying out physical activities and their knowledge on the impact of physically active or inactive towards health
d)	Utilisation of modern medicines and traditional treatment	i.	Identify contributing health beliefs and behaviour for modern and traditional practices

3.5.7 Data Collection

Data collected through in-depth interviews. The in-depth interviews in qualitative research enabled the respondents to provide their views compared to the possibility of being swept over by group dynamics in focus group discussions (Gill et al., 2008). In-depth interviews were conducted using a semi-structured questionnaire in the local Malay language. The interview guide consisted of several key questions that helped to define the areas explored. It also allowed the interviewer or interviewee to pursue an idea or response in more details (Rowley, 2012). When the interview

guide was utilised, it would ensure all the issues of interest were covered (Appendix 12).

The respondents were recruited in April 2019. The interview took place at the community hall of the Kampung Orang Asli Ulu Kelaka. It was essential to interview the respondents in a place where they felt comfortable (Gill et al., 2008). Appointments scheduled with the respondents and agreements regarding the time and the venue for the interviews made.

The interviews were recorded in the Audio Layer-3 Moving Picture Experts Group (MPEG) known as the MP3 format using the ASUS z002 sound recorder (AsusTek Computer Inc.; Taiwan). It had a high sensitivity, low-noise microphone system that filtered noise and had direct USB connectivity. The device was checked prior to the interview session. The recorder was placed on the table to record the conversation. Each interview lasted between 30 and 45 minutes. In addition to the audiotape recorder, the interview was documented using a field notebook if there were any technical issues with the recorder.

3.5.8 Qualitative Data Analysis

For the qualitative data analysis, the process followed the procedures set by Creswell (2014). The process involved preparing the data for analysis. As stated by Rowley (2012), the process of transcribing should start as soon as after each interview to make sense of the interview data. So, after an interview was carried out, the audio recording was transcribed and organised for data analysis. To assist in transferring and coding, the completed interviews were transcribed verbatim in Microsoft Office Word 2016 (Microsoft Corporation; United State) by the researcher. It was maintained in the original language of the interview, which is in Malay.

After that, the researcher explored the data by reading through the interview transcripts and the researcher's notes. Qualitative data were analysed using a content-based thematic approach guided by the Graneheim & Lundman (2004) frameworks. Qualitative data processing was organised in three rounds. In the first round, a codebook was created based on interview questions and answer codes were inductively developed based on the first 10 participants' answers. In the second round, the first 10 transcripts were duplicated to ensure a coders' consensus, to address any inconsistencies in codes and to extend or modify the meanings of codebooks as required. Finally, the codebook was added to the remaining transcripts, and the research team met to review the quotations for each code, to synthesise the content, and to compare inside and through the transcripts, essential quotations that best represented each theme (Puvenesvary et al., 2008).

Below are several considerations made during the qualitative data analysis process:

i. I conducted all the interview sessions. Furthermore, immediately following each interview, I reflected on the session and I transcribed the recording into English at the soonest possible time. The direct involvement in both interviews and transcriptions allowed me to become immersed in the data and gave me a better understanding of the respondents' perspectives. It was a dynamic process and I continuously refined to improve its appropriateness and flow to enhance data collection.

ii. As the respondents were purposely sampled, the findings were not fully applicable to the Orang Asli community living in Jelebu. Nevertheless, in the Phase Two interviews, the intention was to listen to their voices and to record

their views in greater depth, rather than relying on quantitative numbers and figures to represent their view on NCDs.

iii. It was difficult to obtain male respondents as most of the male Orang Asli were not interested or not available during the arranged interview sessions. Thus, it was possible that the views and opinions among male Orang Asli were missed in this qualitative interview.

3.5.9 Trustworthiness of the Findings

Reliability and validity are critical facets of quantitative analysis, though, in the qualitative analysis, it is known as trustworthiness (Graneheim & Lundman, 2004). A few methods have been utilised in this research to ensure that the conclusions of the study are relevant.

a) Interviewer Training

According to Chua (2014), to ensure the validity of the instrument, the interviewer needs to be objective and have an open mind. They must be able to understand the study data clearly and to attempt to elaborate the data accurately without being influenced by their subjective feelings and views. Therefore, the interviewers must be trained beforehand to ensure that the interviews are conducted with precision and detail. The same researchers involved in conducting the quantitative phase were also involved in completing the qualitative phase. Therefore, a good understanding and rapport were established with respondents before the in-depth interview was conducted.

b) Interview Reliability

The transcription was checked to ensure that it did not contain any major mistake. Transcription quality assurance was applied. The codes and themes generated during data analysis were continually compared with the memos write up for the codes and its definition in the Microsoft Office Word 2016 (Daymon & Holloway, 2010). The codes were also discussed and cross-checked by the supervisors (cross-checking). This is known as an external verification process or member checking (Cohen & Crabtree, 2008). It was done by having the committee supervisor checking on the interpretation of data and transcripts. When disagreements arose regarding the description of the findings, discussions were held until a consensus was reached.

c) The Audit Trail

The audit trail define as the detailed record of the decisions made before and during the research and a description of the research process (Daymon & Holloway, 2010). In this study all the documentation, such as raw data, field notes, data collection and analysis procedures we re keep and recorded, so that you can give evidence from them when necessary. The audit trail begins at the proposal stage of the research where we highlight any explicit decisions you have already taken about the methodological and analytical choices of this study. As we begin to collect data, we made a note of extended transcripts, extensive field-notes notes of reflexive and analytical thoughts, as well as a research database.

As proof of the reliability of the qualitative research, the interview session was recorded using an audio recorder. The audio recording enabled the interviewer to capture and preserve the actual language of the interviewee, thus increasing the

accuracy of the data collection. These all contribute to the dependability of your research.

d) Data translations

Data translations were done for chosen quotations only for the purpose of the analysis of the findings. This was done by a second translator to ensure accuracy (Westgard et al., 2019). The data were then presented in the form of a narrative and quotation obtained from an interview transcription.

e) Triangulation of methods

Methodological triangulation has been applied to determine the reliability and confirmability of the qualitative finding. The triangulation process can be used to describe the corroboration between two sets of results or to describe a process of researching a problem using separate methods to obtain a full picture (O'Cathain, Murphy & Nicholl, 2010) In this study, information from quantitative results was used to support the themes that emerged from qualitative research (Moorley & Cathala, 2019).

3.6 Integration of Data

Integration refers to the phases in the research process, where quantitative and qualitative methods are integrated or mixed (Creswell, 2014). In the explanatory sequential design of mixed-methods, integration may occur when selecting the respondents to be followed up in the qualitative process, designing the protocol for qualitative data collection and discussing the results (Halcomb & Hickman, 2015).

Throughout this study, the convergence between the two phases occurred at the intermediate level, where the preliminary findings of the quantitative process

direct the qualitative phase of data collection. The results helped to develop the framework for the qualitative data collection. Another linking point was the analyses, where analyses for both stages were combined to address the research question and give a deeper meaning to the study findings during the discussion of the study results.

3.7 Summary of the Chapter Three

In summary, this chapter offers a comprehensive overview of the methods used for the data collection and analysis. In the Phase One, systematic review conducted following PRISMA guidelines to ensure the a transparent and complete reporting of this type of research. Then a field research conducted at Jelebu, Negeri Sembilan. The sequential explanatory strategies of mixed-method were the most practical approaches for this study. It is more manageable by dividing the tasks into the quantitative phase followed by the qualitative phase. At the same time, it also providing deepen understanding on the findings from the quantitative phase.The techniques for both the quantitative and qualitative phases were explained separately.

A cross-sectional method had been used in the quantitative process. The research was performed using face to face interview questionnaires in eleven Orang Asli villages in Jelebu, Negeri Sembilan, which had been validated and piloted. In the second phase, data was collected through an in-depth interview to explore the factors identified in the quantitative phase. The convergence of both phases first occurred when the quantitative process was used to inform the respondents requirements to be chosen for the qualitative phase. Next, it directs the development of the interview protocol. Phase One systematic review, Phase Two quantitative results and Phase Three qualitative findings were presented in the following chapter.