RECRUS
Res. Newsl.

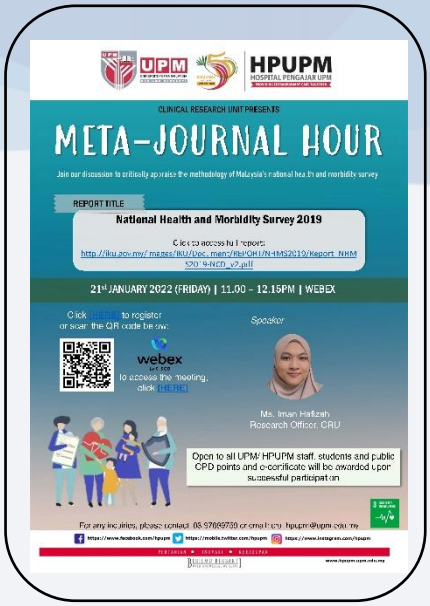
CLINICAL EPIDEMIOLOGY

February 2022
Vol. 2 Issues 12
Page 182

APPRAISALS IN META-JOURNAL HOUR 5

By: Nurul Iman Hafizah, Aazifah Ilham and BH Chew

The paper:
National Health and Morbidity Survey 2019

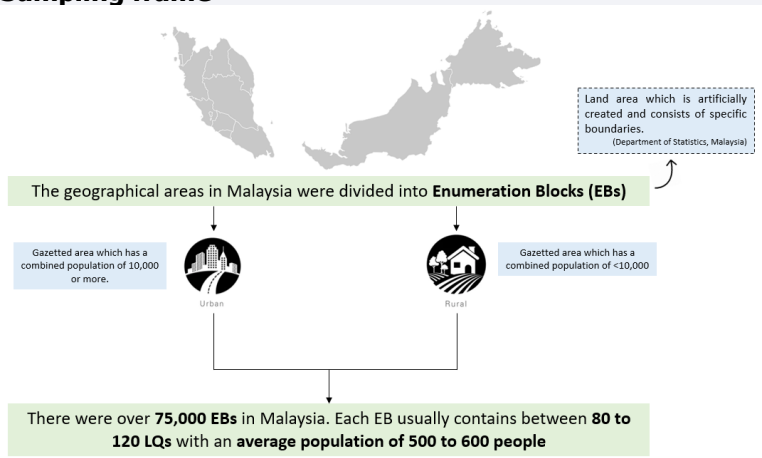


Why was this study conducted?
Non-communicable diseases (NCDs) are the leading cause of mortality worldwide. In Malaysia, NCDs account for 67% of premature mortality, and over 70% of disease burden in 2014. The National Health and Morbidity Survey (NHMS) is carried out to obtain community-based data on the pattern of health problems and health needs of the people in Malaysia. The first NHMS was carried out in 1986

How was it done?

Target population
The NHMS 2019 covered both urban and rural areas in all 13 states and 3 federal territories in Malaysia targeted residence in the non-institutional living quarters (LQs), while those staying in institutional LQs such as hostels and hospitals were excluded from the survey.

Sampling frame



Watch the video recording on:

YouTube

Click [HERE] and don't forget to subscribe to our channel!

Sample size determination
Sample size was calculated using single proportion formula for estimation of prevalence.

$$n_{SRS} \geq \frac{Z^2_{\alpha/2} P(1-P)}{e^2}$$

The sample size calculation was based on a few criteria:

i. Variance of proportion of the variable of interest (Based on NHMS 2015 or other literatures)
ii. Margin of error (e) (Between 0.02 to 0.07)
iii. Confidence Interval of 95%

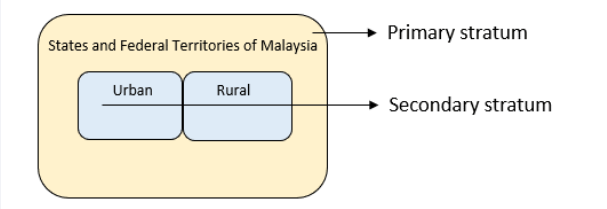
To ensure optimum sample size, few adjustments were made:

- Adjusted for finite population (Based on 2019 projected population)
- Adjusted for the design effect (deff) (Based on previous survey: NHMS 2015)
  n(complex)= n(srs) * deff
- Adjusted the n(complex) taking into account expected non-response rates of 35%
  n(adj) = n(complex) * (1 + non-response rate)

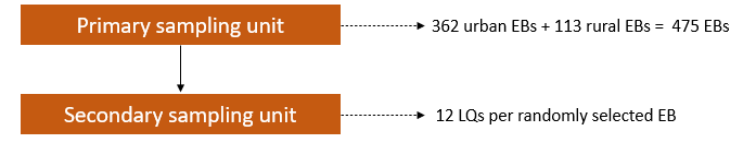
In this survey, sample size calculation was calculated from the variables of interest from the previous NHMS cycle. This also includes the variance of proportion and the margin of error. Smaller margin of error indicates narrow confidence interval and higher confidence level. Small margin of error indicated the sample population closely represents the population of interest. In addition, NHMS also typically involved large number of sample size which also could explain smaller margin of error. From the calculation, the optimum sample size required was 5676 LQs. Bigger number of samples were allocated to states with bigger population size such as Selangor, Johor and Sabah while states with smaller population size such as Melaka, Perlis and Labuan had lesser number of samples allocated.

Sampling design

To ensure national representativeness, two stage stratified random sampling was used.



Sampling involved two stages: the Primary Sampling Unit (PSU), which were the EBs and the Secondary Sampling Unit (SSU) which were the LQs within the selected EBs.



In this study, all individuals with their primary residence and residing for at least 2 weeks prior to data collection in the selected LQs were included in the study.

Questionnaire and Other Survey Materials

Structured questionnaires were used to collect data based on the scopes of the survey. There were two types of questionnaire; face-to-face interview and self-administered. For the face-to-face interview, the pre-tested questionnaire was bi-lingual (Bahasa Melayu and English) accompanied with questionnaire manual prepared as a guide to the data collectors. The self-administered questionnaires were in four languages; Bahasa Melayu, English, Mandarin, and Tamil. There were flash cards of relevant pictures and diagrams provided in the form of code book to assist in the interview. The face-to-face interview questionnaire was programmed into an application and the data collection was done using tablets. Respondents were given the tablet to self-administer the questionnaires. Hardcopies of the self-administered questionnaires were also prepared should the respondent choose to answer in paper. For respondents aged below 13 years, the parent or guardian would respond to the interview on their behalf (by proxy).

Module	Questionnaire	Method	Target Age Group
Household Information	-	Face-to-face	All
Sociodemography	-	Face-to-face	All
Diabetes	STEPS	Face-to-face	18 years and above
Hypertension	STEPS	Face-to-face	18 years and above
Hypercholesterolemia	STEPS	Face-to-face	18 years and above
Physical Activity	IPAQ – Short Form	Face-to-face	16 years and above
Smoking	Mini GATS	Face-to-face	15 years and above
Dietary Practice	-	Face-to-face	18 years and above
Health Screening	-	Face-to-face	18 years and above
Alcohol	AUDIT	Self-Administered	13 years and above
Substance Abuse	-	Self-Administered	18 years and above
Disability	WG Short Set	Face-to-face	18 years and above
Child Functioning	WG	Face-to-face	2 - 17 years
Mental Health (Adult)	PHQ	Self-Administered	18 years and above
Mental Health (Children)	SDQ-Mall	Self-Administered	5 - 15 years
Health Literacy	HLS-M-Q18	Self-Administered	18 years and above
Benign Prostatic Hyperplasia	IPSS	Self-Administered	40 years and above
Erectile Dysfunction	IIEF	Self-Administered	18 years and above
Epilepsy	Ottman Epilepsy Screening	Face-to-face	All

Other assessments such as anthropometry, biochemistry and blood pressure assessments were conducted by a trained nurse.

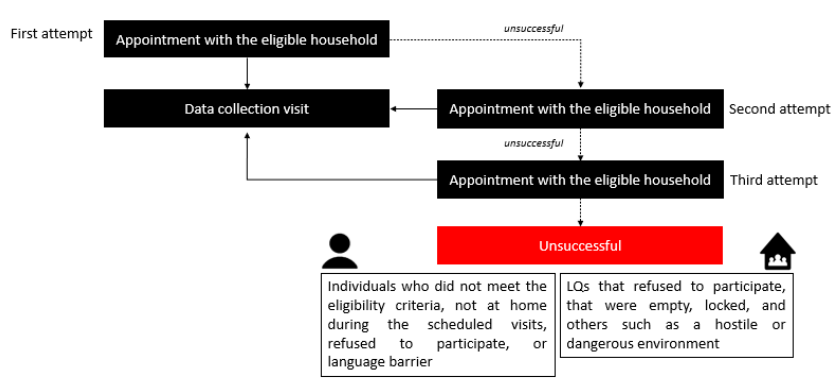
Assessment	Parameters	Population	Tools	Measurer
Anthropometry	Weight	All ages	Tanita Personal Scale HD 319 (adults) Tanita Baby Scale 1583 (infants)	Nurse
	Height/length		SECA stadiometer 213 (adults) Measuring Mat SECA 210 (infants)	
	Waist circumference			
Biochemistry	FBG and cholesterol	≥18 years	CardioChek® PA Analyzer	
	Haemoglobin	≥15years	HemoCue® Machine Hb 201+	
Clinical	Blood pressure	≥18 years	Omron Japan Model HEM-907	

Training

Each team was led by a team leader and comprised of an additional two research assistants (RAs), one nurse and a driver. There was a total of 70 teams, with 57 teams allocated in Peninsular Malaysia and 13 teams in East Malaysia. Training was given to each data collection team to familiarize data collection teams with the questionnaire, develop the interpersonal skills and to appreciate the need for good teamwork. During the training, RAs were briefed on the questionnaire and mock interview practice session under supervision was held. In addition, nurses were trained on the techniques of using the equipment used for clinical assessment and briefed on the criteria for referral of respondents with health problems. At the end of training, pilot test for data collection was conducted.

Data collection

Data collection was carried out between 14th July 2019 to 2nd October 2019.



Data management and monitoring during data collection

- Data processing activities were centralised at the Institute for Public Health (IPH); received data from the field (input from mobile tablet devices to the centralised server) up to handing over the cleaned dataset to the data analysis team.
- Datasets were continuously monitored for quality control; especially on accuracy of the respondent ID, outliers or incorrect data.
- Completed interviews were sent to the Survey Creation System (SCS) server whenever there was an internet connection.

- Quality checks were also in-built into the application pertaining to eligibility in answering different modules based on age group or sex.
- Central Coordinating Team conducted weekly meeting to monitor the progress of each team.
- Productivity of each team was monitored by comparing the cumulative targeted LQs with the weekly progress report by the teams and the amount of data received in the server.

Data analysis

Data analysis was done by the Biostatistics and Data Repository team from the National Institutes of Health, Ministry of Health Malaysia. Analyses were carried out according to objectives of the survey, working definitions and dummy tables. Complex samples analysis procedures were used in the analysis and was carried out at 95% confidence interval. A **weighting factor** was applied to each individual to adjust for non-response and for the varying probabilities of selection.

***In order to make sure that you have a representative sample, you could add a little more "weight" to data.
***Weights are commonly assigned to respondent records in a survey to make the weighted records represent the population of inference as closely as possible
***Weighting factors are used in sampling to make samples match the population.

What was the finding?

Sample coverage

From the 5,147 eligible LQs, a total of 4,703 LQs were successfully interviewed, giving a LQ response rate of 91.4%. From these LQs, a total of 15,683 participants were eligible for interview. A total of 14,965 respondents were successfully interviewed, giving an individual response rate of 95.4%. The overall response rate for this community-based survey is 87.2%.

Sociodemographic characteristics

A total of 10.9% of the respondents of NHMS 2019 are from the state of Selangor. A total of 61.4% of the respondents are from urban localities and females made up 52.8% of the total respondents. Those of Malay ethnicity were 66.5% of the respondents, with Indian ethnicity, Bumiputera Sabah, Bumiputera Sarawak and others making up between 4% to 7% each. A total of 36.5% of the respondents reported their highest educational level to be up to secondary education, with those of no formal education and tertiary education at 16.2% each. Private employees made up 20.2% of the respondents, meanwhile government employees made up 7.0%, with a further 3.4% of retirees in the sample.

Based on the self-reported income of each individual, the household income was calculated and categorized based on state-specific cut-off for B40, M40 and T20 category. The cut- off values for each state were obtained from the Departments of Statistics Malaysia. A total of 68.1% of the respondents fell in the B40 category, with only 8% in the T20 category. This, as in most self-reported studies, shows a certain degree of under-reporting present in the income levels.

Non-communicable diseases

Diabetes mellitus

In this survey, raised blood glucose is defined as fasting blood glucose of ≥7.0mmol/L or non-fasting blood glucose ≥11.1mmol/L. The **overall prevalence** of raised blood glucose among adults aged 18 years and above in this survey was **18.3%** (95% CI: 17.08, 19.58). The prevalence of overall raised blood glucose increased with age, from 5.4% (95% CI: 3.66, 7.91) in the 20-24 years age group, reaching a peak of 43.4% (95% CI: 37.37, 49.65) among the 65-69 years old. The prevalence was reported to be 5.4% (95% CI: 3.66, 7.91) among those aged 20-24 years old and peaked at 65-69 years old [43.4% (95% CI: 37.37, 49.65)]. The prevalence was similar in both urban and rural areas.

The prevalence of **known diabetes** was **9.4%** (95% CI: 8.66, 10.20). The prevalence of known diabetes increased with age. The prevalence was reported to be 4.5% (95% CI: 3.04, 6.71) among those aged 35-39 years and peaked at 65-69 years of age 34.3% (95% CI: 28.59, 40.50). The prevalence in the urban areas was 9.7% (95% CI: 8.83, 10.74) compared to rural [8.2% (95% CI: 7.27, 9.19)].

The prevalence of **raised blood glucose amongst those not known** to have diabetes was **8.9%** (95% CI: 7.96, 9.93). The prevalence of raised blood glucose amongst those not known to have diabetes increased with age. The prevalence was reported to be 4.8% (95% CI: 3.15, 7.33) among those aged 20-24 years old and peaked at 45-49 years old [14.2% (95% CI: 11.17, 17.85)]. The prevalence was higher in the rural, 10.2% (95% CI: 8.66, 11.87), compared to the urban areas,8.5% (95% CI: 7.44, 9.79).

Hypertension
In this survey, raised blood pressure is defined as blood pressure of systolic blood pressure measurement of ≥140 mmHg or diastolic blood pressure of ≥90 mmHg. The **overall prevalence** of hypertension among adults aged 18 years and above in this survey was **30.0%** (95% CI: 28.57, 31.50). The prevalence of overall raised blood pressure increased with age. The prevalence was reported to be 5.7% (95% CI: 4.09, 7.82) among those aged 20 - 24 years old and the highest prevalence of 81.7% (95% CI: 77.03, 85.55) among those 75 years old and above. The prevalence was higher in the rural areas [32.8% (95% CI: 30.03, 35.63) compared to urban areas [29.2% (95% CI: 27.57, 30.97)].

The prevalence of **known hypertension** was **15.9%** (95% CI:14.79, 17.05). The prevalence increased with age, from 1.5% (95% CI: 0.83, 2.51) in the age group 25-29 years old, and the highest prevalence of 65.0% (95% CI: 59.27, 70.23) among the 75 years and above. The prevalence of known hypertension was highest in the rural areas [17.2% (95% CI: 15.37,19.29)] compared to urban areas [15.5% (95% CI: 14.21,16.89)].

Amongst those **not known to have hypertension**, the prevalence of raised blood pressure was **14.1%** (95% CI: 13.13, 15.19). There was a general increasing trend in prevalence with age, from 4.6% (95% CI: 3.17, 6.65) in the 20–24-year-old age group, reaching a peak of 22.6% (95% CI: 18.48, 27.35) among the 45-49 years old. The prevalence was higher in the rural areas [15.5% (95% CI: 13.85, 17.36)] compared to urban areas [13.7% (95% CI: 12.55, 15.00)].

Hypercholesterolemia
In this survey, raised total cholesterol is defined as blood cholesterol level of total cholesterol of 5.2mmol/L or more. The prevalence of **overall raised blood cholesterol** in this survey was **38.1%** (95% CI: 36.15, 40.00). The highest prevalence of overall raised blood cholesterol was among the 70 - 74 years age group [63.4% (95% CI: 56.57, 69.65)].

The prevalence of **known hypercholesterolaemia** was **13.5%** (95% CI: 12.51,14.51). Almost half of the respondents from the 70-74 age group [45.5% (95% CI: 39.02, 52.20)] were aware that they had hypercholesterolaemia.

For this survey, the prevalence of **raised total cholesterol** amongst those not known to have hypercholesterolaemia was **24.6%** (95% CI: 23.03, 26.19). The highest prevalence of raised total cholesterol amongst those not known to have hypercholesterolaemia was recorded among those in the 50-54 years age group.

The detailed findings from NHMS 2019 can be further accessed via this LINK.

How much can we take out from this research/paper?

Proclaiming and broadcasting research findings without in-depth understanding of the methods through which the findings come about is not a practice of responsible persons in scientific research and academia. Such behaviour is near to doing a violence to a respectable scientific and clinical research. It is accepted that there is no flawless clinical studies. If there is one it would become immediately 'flawed' at the time of analysis or results dissemination by the sheer fact that the samples/population changes as the time progresses making the results 'unreal' and 'untrue'. In a more scientific and epidemiological perspectives, sampling and probabilistic analysis come with them the inherent margin-of-error, chance occurrences and statistical assumptions. Knowing and understanding these are keys to appreciating and appraising all clinical research, and when applying the results to life, living and

professional practices. This is because good research requires just a good sampling of subjects and not a census to get to the admissible 'truth', which would in turn be useful to inform and to resolve relevant problems at hand. Therefore, a clearer report on the methodology in NHMS will be informative and educational to those who need it. Similarly, a section on the strength and limitation of the survey will help a better interpretation and application of the results.

One of the important things to bear in mind when reading a research report including NHMS is to grasp a good picture of the study sample characteristics as this would render applicability of the results to your immediate social or professional circles. A separate sociodemography for each outcome measure studied in NHMS will definitely help towards this aspect. On the same note, the ethnic group 'Others' is more appropriately reported in details or where it is feasible or on relevant chapters. Additionally, NHMS is a population-based survey that collect individual-level data. The sampling process is based on EBs and LQs rather than individual persons. This may explain the observed lack of representativeness of the overall samples compared to the Malaysian population at large due to disregards of this individual characteristics in the sampling of LQs instead of individuals. Since the health outcomes measured in the NHMS are individual-level outcomes, many personal and social determinants of health would affect the prevalence of the health indicators. Perhaps, a different sampling strategy that is based on individual characteristics such as age, gender or occupations and then their households could improve on the representativeness issue. This was used to some extend in other similar national healthy surveys [1-5]. This may direct the questionnaires or measures to the right proportion of samples such as men's heath or women's health to the men and women samples, respectively, or the children's outcome to those in the paediatric age group. Inflated sample size of the EBs or LQs might aggravate the disproportions of certain individual-level characteristics such as the ethnic groups and income level seen in the NHMS samples. Inflation of sample size to ensure adequate power at state-level and at the same time to maintain the acceptable proportions of sample between-states by inflating it in other states might be a costlier approach than the sampling based on the individual characteristics. However, as explained by the PI this issues could have been minimised by the statistical weighting in the analysis. A clearer description on this statistical strategy will educate and help readers to comprehend and be convinced, as well as for the investigators and authors in better reporting and results discussion. Weighting strategy is like other statistical strategies ie multiple imputation for missing data, that could introduce artificial value if it is not properly conducted and verified.

In summary, I believe the results from NHMS 2019 are valuable health data of the Malaysian population to-date. As shared by the PI Dr. Shubash Shander the challenge in getting responses from the identified population is a real, and it is truly an amicable feat in conducting it on the ground. Perhaps, the time has arrived that the next NHMS to consider a different approach that is more efficient and effective that take into consideration the life in the endemic of COVID-19, taking advantage of high penetration of smart mobile devices in the population, altered health behaviours and awareness [1-9]. This would hopefully also improve the variables chosen and the responses given to them are of credible value, either they come from self-reporting or linking up to personal data from auto-generated administrative databases. Having an independent data monitoring committee to verify the study process and data will provide NHMS an international reputation and quality.

References

1. Ministry of Health. 2021. Methodology Report 2020/21: New Zealand Health Survey. Wellington: Ministry of Health. Available from: https://www.health.govt.nz/system/files/documents/publications/methodology-report-2020-21-new-zealand-health-survey-dec21.pdf
2. Statistics Canada. National Population Health Survey: Household Component, Longitudinal (NPHS). Available from: https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3225&lang=en&db=imdb&adm=8&dis=2
3. National Health Survey. Australian Bureau of Statistics. Available from: https://www.abs.gov.au/methodologies/national-health-survey-first-results-methodology/2017-18
4. European Health Interview Survey (EHIS). https://ec.europa.eu/eurostat/cache/metadata/en/hlth_det_esms.htm
5. National Research Council (US) Committee on National Monitoring of Human Tissues. Monitoring Human Tissues for Toxic Substances. Washington (DC): National Academies Press (US); 1991. Appendix D, The National Health and Nutrition Examination Survey. Available from: https://www.ncbi.nlm.nih.gov/books/NBK234172/
6. National Research Council (US) Coordinating Committee on Evaluation of Food Consumption Surveys. National Survey Data on Food Consumption: Uses and Recommendations. Washington (DC): National Academies Press (US); 1984. 3, THE NATIONAL HEALTH AND NUTRITION EXAMINATION SURVEY. Available from: https://www.ncbi.nlm.nih.gov/books/NBK217716/
7. Picavet HS. National health surveys by mail or home interview: effects on response. J Epidemiol Community Health. 2001 Jun;55(6):408-13. doi: 10.1136/jech.55.6.408
8. Fowler FJ Jr, Gallagher PM, Stringfellow VL, et al. Using telephone interviews to reduce nonresponse bias to mail surveys of health plan members. Med Care. 2002 Mar;40(3):190-200. doi: 10.1097/00005650-200203000-00003
9. Mlikotic R, Parker B, Rajapakshe R. Assessing the Effects of Participant Preference and Demographics in the Usage of Web-based Survey Questionnaires by Women Attending Screening Mammography in British Columbia. J Med Internet Res. 2016 Mar 22;18(3):e70. doi: 10.2196/jmir.5068
10. Shim JM, Shin E, Johnson TP. Self-rated health assessed by web versus mail modes in a mixed mode survey: the digital divide effect and the genuine survey mode effect. Med Care. 2013 Sep;51(9):774-81. doi: 10.1097/MLR.0b013e31829a4f92