Research and Reports in Tropical Medicine

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

The quality of sample surveys in a developing nation

This article was published in the following Dove Press journal:
Research and Reports in Tropical Medicine
3 September 2010
Number of times this article has been viewed

Paul A Bourne[1]
Christopher AD Charles[2,3]
Neva South-Bourne[4]
Chloe Morris[1]
Denise Eldemire-Shearer[1]
Maureen D Kerr-Campbell[5]

[1]Department of Community Health and Psychiatry, Faculty of Medical Sciences, University of the West Indies, Mona, Kingston, Jamaica; [2]King Graduate School, Monroe College, Bronx, New York, USA; [3]Center for Victim Support, Harlem Hospital Center, New York, USA; [4]Research assistant for Paul A Bourne; [5]Systems Development Unit, Main Library, Faculty of Humanities and Education, University of the West Indies, Mona, Kingston, Jamaica

Background: In Jamaica, population census began in 1844, and many intercensal ratios obtained from the census data showed that there is a general high degree of accuracy of the data. However, statistics from the Jamaican Ministry of Health showed that there are inaccuracies in health data collected from males using sample surveys.

Objectives: The objectives of the present research are to 1) investigate the accuracy of a national sample survey, 2) explore the feasibility and quality of using a subnational sample survey to represent a national survey, 3) aid other scholars in understanding the probability of using national sample surveys and subnational sample surveys, 4) assess older men's evaluation of their health status, and 5) determine whether dichotomization changes self-evaluated health status.

Methods: For the current study, the data used in the analysis were originally collected from 2 different sources: 1) the Jamaica Survey of Living Conditions (JSLC) and 2) Survey of Older Men (SOM). Cross validation of self-evaluated data of men in Jamaica was done with comparable samples of the complete JSLC data and the SOM data, where men older than 55 years were selected from each sample.

Results: In study 1, 50.2% of respondents indicated at least good self-evaluated health status compared with 74.0% in study 2. Statistical associations were found between health status and survey sample (χ^2 [df = 5] = 380.34, $P < 0.001$), self-reported illness and study sample (χ^2 [df = 1] = 65.84, $P < 0.01$, $\phi = 0.16$), and health care – seeking behavior and study samples (χ^2 [df = 1] = 21.83, $P < 0.05$, $\phi = 0.10$). Substantially more respondents reported an illness in study 1 (34.3%) than in study 2 (ie, 17.5%).

Conclusion: Clearly, inconsistencies exist in the health data, which indicates that care should be taken in using sample surveys.

Keywords: validation, health data, males, public health, national surveys, subnational surveys

Introduction

This article examines the accuracy of a national survey, assesses the usefulness of using a subnational sample survey to understand the national survey, and attempts to act as a guide to fellow researchers. This article used self-evaluated data from older men on their health status and seeks to elucidate whether self-evaluated health status changes with the dichotomization process. Since 1844, census-taking has been an irregular decennial event in Jamaica (with none done in the 1930s). Statisticians and researchers have obtained many intercensal ratios from the data, which showed a high degree of accuracy.[1] This occurrence is also the case in North America, Britain, Japan, India, Africa, and Europe.[1] Vital registration statistics (data on births, deaths, and

Correspondence: Paul A Bourne
Research Fellow and Biostatistician, Department of Community Health and Psychiatry, University of the West Indies, Mona, Kingston, Jamaica
Tel +1 876 457 6990
Email paulbourne1@yahoo.com

marriages) have been used for years in the computation of life expectancies, health, and prosperity of nations. Census-taking and civil registrations are highly expensive data collection processes, which accounts for the use of sample surveys. The first national sample survey was in 1953 to aid intercensal planning. The use of survey data by nations in planning denotes that planners in particular researchers rely on the completeness and accuracy of the data.

Sample surveys are widely utilized to examine social conditions. They are also used for much more than the understanding of social conditions, which include life expectancy, mortality patterns, fertility, termination of marriage, population projections, other demographic computations, and health statistics.[1–6] Unlike a census, a survey collects standardized data from a specific population with the purpose of generalizing this to a wider population.[7] In the sampling and data collection processes, errors are highly likely to enter into the data.

The quality of sample survey data is important for more than the accuracy of using sampling design in a particular task. The guidance that sample survey methods provide to researchers is embedded within people. It follows that sample survey must rely on the accuracy in the recall and truth of information provided by research participants. This information not only influences people socially but also impacts on the quality and quantity of their lives. It is within this context that the accuracy of sample surveys is crucial to researchers, policy makers, and nonacademics, as they seek to enhance the quality and quantity of human experience. This mindfulness requires that researchers take into account the broadest possible range of reasons within the parameters of the research that brings validity and reliability into disrepute.

Jamaica is among many countries that collect sample survey data to guide, formulate, assess, and understand its population. In 1988, Jamaica began collecting sample survey data on the living standard of its people. The survey is referred to as the Jamaica Survey of Living Conditions (JSLC). The JSLC is a modification of the World Bank's Living Standards Measurement Study (LSMS), which is a household survey. The JSLC provides policy makers including the government with vital information on the policy implementations and their effect on the living standard of people. Health, which is more than disease,[8] means that JSLC coverage is comprehensive enough to allow for the assessment of health of Jamaicans.

Using Jamaican Ministry of Health Annual Reports on the actual visits made to health care facilities as well as visits for curative care, Table 1 shows that on average 30% of males visited health care facilities and 34% received curative care. However, survey statistics for the same period showed that on average healthcare visits for males was 62% and self-reported illness was 10%.[9] This highlights inconsistencies in the data sources. Within the context of disparities that exist in the data sources, this brings into question the reliability and validity of health survey data, which are collected from males in Jamaica.

Table 1 Health and curative care visits: 2000–2007

Year	Health care visits			Curative care visits		
	Male	Female	Female:male	Male	Female	Female:male
2000	30.5	69.5	2.3:1	35.0	65.0	1.9:1
2001	30.6	69.4	2.3:1	34.6	65.4	1.9:1
2002	30.3	69.7	2.3:1	34.2	65.8	1.9:1
2003	30.3	69.7	2.3:1	33.5	66.5	2.0:1
2004	30.2	69.8	2.3:1	32.9	67.1	2.0:1
2005	30.4	69.6	2.3:1	33.3	66.7	2.0:1
2006	30.4	69.6	2.3:1	33.5	66.5	2.0:1
2007[a]	30.5	69.5	2.3:1	33.6	66.4	2.0:1

Note: Figures were computed by Paul A Bourne from Jamaica, Ministry of Health (Jamaica) Annual Report 2004 and 2006. [a]Preliminary data from the Ministry of Health (Jamaica) were used to compute the percentages and ratios.

A critical assessment of the literature has revealed that there is a paucity of research investigating the validity of sample survey data in the Caribbean in general and Jamaica in particular. The Caribbean, like many other nations, has come to rely on sample survey data in government and health planning. People's lives, therefore, cannot be based on inaccuracies from sample survey data, and so an examination of the accuracy of survey will allay many fears of critical stakeholders and nonresearchers. Validity is critical for understanding and interpreting already published studies and to guide new studies and survey approaches. The objective measurements are infrequently used to validate costly questionnaires, such as the JSLC. Agreement reality trumps reliability and validity where without question the data are taken as accurate over time and that the instrument is measuring what it purports to measure.

The current study examines the accuracy of 2007 JSLC by using another independent sample survey in the same period. The objectives of the present research are to 1) investigate the accuracy of a national sample survey, 2) explore the feasibility and quality of using a subnational sample survey to represent a national survey, 3) aid other scholars in understanding the probability of using national sample surveys and subnational sample surveys, 4) assess older men's evaluation of their health status, and 5) determine whether dichotomization changes self-evaluated health status.

Methods and materials

Data

For the current study, the data used in the analysis were originally collected in 2007 from two different studies: 1) JSLC and 2) SOM. In order to cross validate self-evaluated data from men in Jamaica, because complete data were available from JSLC and only data on older men (>55 years) from SOM, participants older than 55 years were selected from each sample, as this was comparable in both samples. For the current study, 2,483 participants were used: 483 participants in study 1 (ie, JSLC) and 2,000 participants in study 2 (ie, SOM). The mean age in study 1 was 67.7 years (SD = 9.3 years), and in study 2, it was 67.0 years (SD = 8.2 years). Urban dwellers comprised 47.0% (n = 227) in study 1 and 49.1% in study 2 (n = 981) compared with 53.0% and 50.9% in rural dwellers, respectively.

Sample

Study 1

Data from the JSLC for 2007 commissioned by the Planning Institute of Jamaica (PIOJ) and the Statistical Institute of Jamaica (STATIN) were used to provide the analyses for this study.[9] These two organizations are responsible for planning, data collection, and developing policy guidelines for Jamaica and have been conducting the JSLC annually since 1989. The cross-sectional survey was conducted between May and August 2007 from the 14 parishes across Jamaica and included 6,782 people of all ages.[10] The sample for this study comprised 1,343 respondents who were classified as in the poorest 20% in Jamaica (or the poorest 20%).

The JSLC used stratified random probability sampling technique that was drawn to the original sample of respondents, with a nonresponse rate of 26.2%. The JSLC survey was based on a complex design with multiple stratifications to ensure that it represents the population, marital status, area of residence, and social class. The sample was weighted to reflect the population.

The instrument used by the JSLC was an administered questionnaire where respondents are asked to recall detailed information on particular activities. The questionnaire was modeled from the World Bank's LSMS household survey. There are some modifications to the LSMS, as JSLC is more focused on policy impacts. The questionnaire covers demographic variables, health, immunization of children aged 0–59 months, education, daily expenses, nonfood consumption expenditure, housing conditions, inventory of durable goods, and social assistance. Interviewers were trained to collect data from household members.

Study 2

The study used primary cross-sectional survey data on men older than 55 years from the parish of St Catherine in 2007 (May and June); it is also generalizable to the island.[11–13] The survey was submitted and approved by the University of the West Indies Medical Faculty's Ethics Committee. Stratified multistage probability sampling technique was used to draw the sample (2,000 respondents). A 132-item questionnaire was used to collect the data. The questionnaire was subdivided into general demographic profile of the sample, past and current health status, health-seeking behavior, retirement status, social and functional status.

The STATIN maintains a list of enumeration districts (EDs) or census tracts. The parish of St Catherine is divided into a number of constituencies made up of a number of EDs. The sampling frame was provided by the 162 EDs in the parish of St Catherine. The EDs were listed and numbered sequentially, and selection of clusters was arrived at by the use of a sampling interval. Subsequently, 40 EDs (clusters) were selected, with the probability of selection being proportional to population size (Table 2).

The sample population does not only speak of the parish of St Catherine, it is generalizable to the island of Jamaica. The sampling frame was men older than 55 years in the parish of St Catherine. The parish of St Catherine was chosen because previous data and surveys[11–13] suggested that it has the mix of demographic characteristics (urban, rural, and age composition) that typify Jamaica.

EDs consisted of not more than 400 households, and they were used as primary sampling units. Interviewers were trained by the staffers of University of the West Indies (ie, Department of Community Health and Psychiatry), and large groups were subdivided into smaller group with a supervisor who monitored the interviewers in an effort to maintain accuracy. All interviewers were given a map of their respective EDs and they were taken across the geographic boundary of that EDs. Stratified random sampling was

Table 2 Proportion of survey (sample) vs proportion of population

Age group, y	Survey		2001 census (St Catherine)		2001 census (Jamaica)	
	n	%	n	%	n	%
55–59	469	23.45	6,577	26.7	38,645	23.9
60–64	413	20.6	5,179	21.1	31,828	19.7
65–69	374	18.7	4,391	17.8	28,901	17.9
70–74	345	17.2	3,594	14.6	24,856	15.4
75–79	189	9.45	2,402	9.78	17,711	11.0
>80	210	10.5	2,399	9.77	19,552	12.1

used to predetermine that who should be interviewed from particular households. Enumerators commenced at a fixed point as was stipulated by STATIN, and the interviewers proceeded based on their map of the predetermined persons in a clockwise manner. This approach was used in order to exhaust the EDs.

In the event a chosen participant from a household did not wish to participate in the interview, the interviewer would go to the next identified household on his or her map. For males older than 55 years who were not at home while the interview was conducted, a maximum of three call backs were used in order to establish link for a possible interview. In cases where the interviewer had exhausted all the call backs and the participants were still unavailable, a replacement was used from the adjacent household assuming that the person fulfils the criterion of the study (ie, males older than 55 years). A strict definition of the household was used as a measure of standardizing those who should be interviewed for the study. A household was any individual who slept in the dwelling for at least three nights and ate at least one meal per week from the same pot as other individuals. Hence, a resident for selection had to be male older than 55 years who slept in the same dwelling as other individuals for at least three nights and ate from the same pot at least once per week as other individuals of that dwelling.

Validity

The current study validated the self-evaluated health data used in the JSLC by using subnational sample survey, which was collected during the same time as that of the former survey. The JSLC questionnaire could not be assessed, as it would be expensive to do so, and the objective of the latter study sought to examine the health status, health literacy, health decisions, and typology of dysfunctions that older men have. Both studies used a 5-point self-rated health status question (Generally, how would you rate your current health status?). The responses ranged from excellent (ie, very good) to very poor, and this allowed for the validation of health status.

Reliability

One of the purposes of using matching studies is to examine content errors that assess the reliability of the data sources.[1–4] Testing the consistency of information derived from the National Survey (ie, JSLC) will be done using a SOM (older than 55 years). The older-men study was conducted in St Catherine and this will be used to assess the consistency of the information in the National Survey. Consistency (or inconsistency) was evaluated by using Chi-square analysis. If there was no association between the variable and the different samples, then the information in one survey was consistent with the information of the other survey. Conversely, statistical association denotes inconsistency of results.

Ethics

No ethical clearance was sought for JSLC. However, one was sought and obtained for the subnational sample survey. The University of the West Indies Ethical Board approved the subnational sample survey, and each participant was given a written informed consent prior to his or her participation.

Statistical analysis

Data were stored, retrieved, and processed using SPSS for Windows 16.0, and a 5% level of significance was used to test significance (ie, 95% confidence interval [CI]). Descriptive statistics were used to provide background information on the samples. Validity was assessed by comparing levels of self-evaluated data on 1) area of residence, 2) age group,(3) health status, 4) marital status, 5) household heads, 6) medical care-seeking behavior, and 7) self-diagnosed illnesses. The researchers also used χ^2 to measure associations between the 2 sample survey results. A statistical association from a χ^2 result should be interpreted as differences between the study 1 (ie, National sample survey) and study 2 (ie, St Catherine sample survey). On the other hand, no relationship should be interpreted as any difference between the aforementioned study samples. Contingency coefficient and χ^2 were used to examine the statistical association between variables.

Measurement of variables

Self-rated health status is measured using people's evaluation of their overall health status,[14] which ranges from excellent to poor. The question that was asked in the survey was "How is your health in general?" And the options were very good, good, fair, poor, and very poor. For the purpose of the model in this study, self-rated health was coded as a binary variable (1 = good and fair 0 = otherwise; also see studies that have treated self-rated health status as a binary variable).[15–20] Age is a continuous variable, which is the number of years alive since birth (using last birthday). For the present study, ages range from 55 years and older.

Data errors for this work are classified into two groups: coverage errors and content errors. Coverage errors arise due to completeness of inclusion of people in a data system.[2,4]

This includes misplacement of events in a time or when events are incorrectly classified in one defined boundary when it should have been an estimate in another defined unit. Content errors denote accuracy in characterization of recorded units in a data system.[2,4] Sampling errors denote negative errors of failure to include elements that should properly belong to a sample against a population. These arise owing to coverage errors. Nonsampling errors are all errors that are theoretically outside of those caused by sample against a population. These include nonresponse, content errors, and interviewers' biases.

Results

Demographic characteristic of sample

The sample was 2,483 respondents (483 for the national survey [study 1] and 2,000 for the subnational survey [study 2]). In study 1, the mean age was 67.7 years (SD = 9.3 years), while the mean age in study 2 was 67.0 years (SD = 8.2 years). In study 1, 50.2% of respondents indicated at least good self-evaluated health status compared with 74.0% in study 2 (Table 3).

In study 1, 34.3% of sample indicated an illness compared with 17.5% in study 2. The percentage of respondents who indicated having sought more medical care was also more in study 1 (ie, 65.5%) than in study 2 (ie, 45.7% [Table 4]).

Bivariate analyses

There was no association between area of residence and study used (χ^2 [df = 1] = 0.66, $P > 0.05$). This was also the case for age and study sample (χ^2 [df = 5] = 8.66, $P > 0.25$). However, relationships were found between (1) marital status and study sample (χ^2 [df = 4] = 15.38, $P < 0.01$, contingency coefficient (C) = 0.08), and (2) household head and study sample (χ^2 [df = 1] = 33.71, $P < 0.01$, $\phi = 0.12$). Furthermore, for study 1, 78% of respondents indicated that they were heads of their households compared with 88% of those in study 2. With regard to marital status and study sample, 10% of those in study 1 revealed that they were in common-law unions compared with 7% of those in study 2. Similarly, percentage point disparity was found in separated unions (ie, 3% in study 1 and 6% in study 2).

A cross-tabulation between self-evaluated health status and study sample revealed a statistical association (χ^2 [df = 5] = 380.34, $P < 0.001$). The relationship between the two variables was weak (C = 0.37 or 37%). Substantially more respondents in study 2 indicated at least good health compared with those in study 1. The converse was equally true as more people in study 1 reported at least poor health compared with those in study 1. When self-evaluated health status was dichotomized (ie, good vs poor health [with moderate or fair health being included in poor health]), the relationship between the self-evaluated health status and the study sample became weaker (ie, $\phi = 0.21$ or 21%; χ^2 [df = 1] = 105.68, $P < 0.05$) than when health status was nondichotomized. When self-evaluated health status was dichotomized, 49.8% of respondents in study 1 indicated moderate-to-very poor health status compared with 26% of those in study 2.

A statistical relationship was found between medical care-seeking behavior and study samples (χ^2 [df = 1] = 21.83,

Table 3 Descriptive characteristic of samples: subnational and national surveys

Characteristic	Subnational survey (study 2)		National survey (study 1)	
	n = 2,000	%	n = 483	%
Area of residence				
Urban	981	49.1	227	47.0
Rural	1,019	51.0	256	53.0
Age group (y)				
55–59	469	23.5	120	24.8
60–64	413	20.7	87	18.0
65–69	374	18.7	88	18.2
70–74	345	17.3	68	14.1
75–79	189	9.5	61	12.6
>80	210	10.5	59	12.2
Marital status				
Single	686	34.3	150	31.8
Married	894	44.7	217	46.1
Separated	112	5.6	14	3.0
Common-law	136	6.8	49	10.4
Widowed	172	8.6	42	8.7
Household head				
Yes	1,763	88.2	376	77.8
No	237	11.8	107	22.2
Self-rated health status				
Excellent (or very good)	357	19.0	61	12.9
Good	1,038	55.4	177	37.3
Fair (or moderate)	480	25.6	149	31.4
Poor	0	0.0	75	15.8
Very poor	0	0.0	12	2.5
Self-evaluated diagnosed illness				
Cold	–	–	14	8.5
Asthma	5	0.3	8	4.8
Diabetes mellitus	129	6.5	24	4.8
Hypertension	193	9.2	24	14.5
Arthritis	20	1.0	46	27.9
Diarrhea	–	–	2	1.2
Other: unspecified	–	–	22	13.3
Other: cancer	336	16.8		
Heart disease	106	5.3		
Kidney/bladder	118	5.9		
Prostate problem	143	7.2		

Table 4 Characteristic of samples: subnational and national surveys

Characteristic	Subnational survey (study 2)		National survey (study 1)	
	n = 2,000	%	n = 483	%
Medical care-seeking behavior				
Yes	914	45.7	106	65.5
No	1,086	54.3	58	34.4
Sought medical care				
≤12 mo	289	31.6	NS	NS
12–35 mo	356	38.9	NS	NS
≥36 mo	269	29.4	NS	NS
Provision of care				
Home remedy	155	44.3	66	13.7
Public clinic	124	35.4	73	15.2
Hospitals	40	11.4	149	30.8
Private doctor	31	8.9	195	40.3
Self-evaluated illness				
Yes	350	17.5	162	34.3
No	1,650	82.5	310	65.7

Note: NS, Not stated (ie, not collected in this study).

$P < 0.05$, $\phi = 0.10$). In study 1, 65% of respondents claimed to have seeked medical care compared with 46% in study 2. Likewise, an association was found between self-reported illness and study sample (χ^2 [df = 1] = 65.84, $P < 0.01$, $\phi = 0.16$). Substantially more respondents reported an illness in study 1 (34.3%) than in study 2 (ie, 17.5%).

Discussion

Accuracy of national surveys and subsample surveys

Scholars and the STATIN have found and purported that stratified sampling of the parish of St Catherine is a good measure as a representative sample of the wider Jamaican population.[1–13] They found that St Catherine has a diverse population with congruent characteristics similar to the wider Jamaican population. The current study found that although a sample of older men in St Catherine had similar characteristics like a National sample (JSLC), some disparities still exist between the two samples. If a sample of St Catherine is similar to that of Jamaica, then there must not be any disparity in medical expenditure, health status, marital status, household head, and self-reported illness. In fact, this study found that even among some demographic characteristics like household head and marital status differences were there between both surveys.

The nonvalidation in some of the demographic characteristics for both surveys was also found in the self-evaluated health data. The current research found that there was a difference between self-evaluated health data for the St Catherine and the National samples. In the St Catherine sample, none of the respondents indicated having poor health compared with 18.4% of those in the National sample. The disparity was more so in the category of good health. In the National sample, 37.3% revealed having good health compared with 55.4% in the St Catherine cohort. This denotes that 1.5 times more respondents indicated that they had good health in latter sample, suggesting that there is either overstatement or understatement of health status among older men in Jamaica. While the current study does not accept that any of the two samples is correct over the other, it is evident from the significant inconsistencies between both the samples that health data from older men are incorrectly reported. Embedded in the health data from older men, therefore, is nonsampling errors[4,21–24] from a finite population, suggesting that public health planning with inaccurate health data will yield low quality health outcomes.

Quality of national surveys and subsample surveys, and health status of older men

Among nonsampling errors is attitudinal information.[1–5] Demographers like Colin Newell[4] believed that despite the extensive training that may be had by sample data collectors, their attitude including appearance can affect the quality of the information that they receive (or not receive) from the interviewee. Caribbean societies in particular Jamaica have not been examining the quality of data collected owing to attitudinal biases. Many of the data collectors in Jamaica are females and with the context that males do not want to appear weak, effeminate, or sick, males reporting illnesses to females clearly are distorting the quality of the data. One Caribbean anthropologist argued that Caribbean males are socialized to be tough, strong, and display no signs of weakness.[25] Another Caribbean scholar opined that sickness is interpreted by Caribbean males as a signal of weakness,[26] which justifies their reluctance to openly speak about illnesses. Males' unwillingness to speak about illnesses crosses gender types, as they must preserve their masculinity both among other males and females.

Some non-Caribbean researchers found that only 10.5% of men who suffer from erectile dysfunctions sought medical care,[27] suggesting that males' have a preference not to speak about and display signs of weakness. Illness, which is an indicator of weakness for males, means that health care – seeking behavior is usually a last resort and is most often used in case of severity of illness.[28] Statistics from the Ministry of Health (Jamaica) showed that on average 30 of every 100 males sought medical care, which denotes that older males were

substantially under-reporting their illness in the St Catherine study. Although 34 of every 100 older males sought medical care, as indicated by the National survey, within the context that there is a positive relationship with aging and poor health status, it can be extrapolated and projected from the Ministry of Health (Jamaica) data[29,30] that this group should not have indicated only 4% more dysfunctions compared with the general population. Few conditions such as common cold and stomach upset are not included in the Jamaicans' definition of illness, as they are treated by home remedy. Of Jamaicans who reported an illness, statistics showed that 30.2% utilized home remedy compared with 28.4% of males and 37.8% of females,[9] which highlights the role of culture in defining and changing health care seeking in Jamaica.

If illnesses do not disrupt males' economic livelihood, many of them are highly unlikely to seek health care because they do not see the need to attend traditional medical facilities as the illness can be rectified at home as they are not ill enough. This is not atypical to Jamaica, as in Pakistan[27] young males were more likely to seek medical care if illness interferes with their economic livelihood. This explains why many males in Jamaica on average spent more time receiving medical care than females,[9] and accounts for the higher mortality,[12] as the time at which they visit health care institutions the dysfunction would have increased to chronic, untreatable and incurable, thus making it highly unlikely for medical practitioners to make a difference. The culture, therefore, retards many Caribbean and non-Caribbean males from truthfully reporting health matters, and the fact that females are collecting information from them about health matters further accounts for the increased non-sampling errors (ie, inaccuracies in data as the under-stated dysfunctions to create an impression of strength, which is tied to their perception of health).

The findings from the present research revealed an exponential disparity between self-evaluated illnesses of the two samples. Approximately twice as many older men reported illnesses in the National survey compared with the St Catherine survey. Therefore, there is a difficulty in validating self-evaluated health data collected from older men in Jamaica. Statistics from the Ministry of Health (Jamaica) showed that 34 of every 100 males received curative care and with the same number from the national survey, it follows that in the St Catherine study there were substantially more under-reported illness.

Although we can extrapolate an exact value for older males, the number of older males who received curative care varies because ageing is associated with increased illness and reduced function, and therefore, the researchers suggest that a greater percent of older males should have reported an illness that was higher than the national average. The reality of the unreliability and invalidity of health data is further highlighted in the self-reported typology of health conditions between the two sample surveys. In both sample surveys, there was not consensus on the typology of dysfunction. In some cases, the disparities were huge as it was for arthritic cases. One percent of respondents indicated that they had arthritis in the St Catherine sample compared with 28% in the national sample.

It should be noted that statistics from the Ministry of Health (Jamaica) between 2002 and 2006[29,30] showed that females received two times more curative care than males. However, self-reported data from the PIOJ, and the STATIN showed that 1.5 times was the greatest disparity, with females reporting more illnesses, and this was 1.4 times more in 2007 (17.8% of females to 13.1% of males). On the other hand, statistics from the Ministry of Health (Jamaica) on curative visits showed that since 2000 on average 34% of males received care.[29,30] This further reinforces the inaccuracies in self-evaluated health data provided by males which the case is using elderly samples for two different sources over the same period. Emerging from the study is that the inaccuracies are not limited to older men but that this is generalizable to the populace of males in nation.

Another area in which disparity was found is medical care-seeking behavior. In the National survey, 66% of older men reported that they visited health care practitioners compared with 46% of sample in the St Catherine study. For 2000 and 2006, statistics from the Ministry of health (Jamaica) showed that approximately 30% of males have been visiting health care agencies. Aging has increased visits to medical care seeking facilities but the figure of 66% of older men seeking medical care as revealed by the National survey seems rather high within the context of already discussed socialization. The discrepancy maybe due to the participants' belief that they should give government agencies the data they want rather than data that is correct. This further brings into question the quality of self-evaluated health data collected from males in Jamaica, and how future studies must be interpreted, ergo, they must incorporate the findings of the current study in their analyses.

Dichotomization of self-evaluated health status

In the validation process of the health data what emerged is the loss of some of the original information from dichotomized self-evaluated health status. Using nondichotomized

self-evaluated health status, the relationship between self-evaluated health status and study samples was 37%, and when self-evaluated health status was dichotomized the association fell to 21%. This concurs with studies that found that it is better to use the continuous nature of self-evaluated health status than the dichotomized variable[31–33] as in the dichotomization process some of the original information will be lost. The current research showed that 16% of the original information is lost owing to the dichotomization process. This highlights a rationale for the nondichotomization of self-evaluated health status in Jamaica as data losses denote the lowering of the quality of health data and foster challenges in policy implementation from a dichotomized health status.

Some studies have found that dichotomizing self-rated health and using logistic regression are acceptable,[16–18] and many studies in Jamaica have followed this procedure[36–41] but clearly using this operational definition in examining the health of males will not produce the same interpretation as some of the original information would have been lost. Recently, a study by Bourne[41] found that dichotomizing self-rated health is acceptable for females, as there was minimal variance; however, much variance was found in dichotomizing health for males. Another study found that when poor self-rated health status was narrowly defined (excluding moderate health), there were minimal impact on the estimated effects of the covariates,[15] and this was reinforced by Bourne's work.[41] However, Bourne's findings somewhat disagreed with the findings of Finnas et al[15] as they show substantial disparity for males when moderate health is included to poor health status.

Validity and reliability of using national surveys and subsample surveys

Inaccuracies are found in the present study as already outlined, but these exclude errors associated with coverage and content. The national survey (ie, JSLC) undoubtedly used complex statistical techniques to design its sample and has reduced coverage errors. The JSLC updated its sample frame in 2007,[9] which adds to the quality of coverage and further reduced coverage errors, as more people would be included in the sample, in order to be better able to select a sample that is more representative of the population. By widening the sampling frame, negative errors of failure are reduced because more elements that belong to the population will be included and thereby can be selected for the national survey sample.[42] However, the quality of the sample coverage does not mitigate against content errors, which appears to be present in the health data. Content error also plays a role in influencing sample outcomes and thereby the quality of data collection.[1–4,21–24] Content errors are a part of response errors and so cannot be neglected in the sampling process, as they act jointly with coverage errors to lower the quality of data collection.[42] It can be said that the inaccuracies found in the health data of older men explain the content errors in the data and cannot be neglected, as these will influence the outcomes and the interpretation of those outcomes. This is also a public health challenge, as not having quality data denotes that policies will address inaccuracies and will further retard all forms of development in the nation.

Surveys on health are among the epidemiological studies executed, and they provide critical information on various health issues, such as dysfunctions, duration of illness, hospitalization, self-rated health, among other variables. Validity of data assists with understanding the quality of health data, and this is agreed with by many demographers[1–4] and nondemographers in the Caribbean.[41] Wilks et al[43] examined the validity of nonresponse and concretized a position that quality health data are based on precision in sample size, nonresponse, and the current study goes further to show that content errors will affect the outcome of the collected data. Interestingly, in addition to Wilks et al many other researchers also embarked on sampling errors, but avoided the importance of content quality. Examining nonresponse errors assumes that content errors are nonexistent or minimal, and even in Wilks et al's work, within the context of the current findings, there are content errors because the study collected data from males who were less likely to report quality information on their health status.

Empirical studies have established that the quality of data in developing countries is relatively low.[1–4,44] Jamaica is a middle-income developing nation in the Caribbean, with a history of high-quality statistical data sources. Using intercensal survey and demographic ratios, it has been shown that the data collected in the censuses and the JSLC are of high quality. The longstanding nature of data collection and the continuous updating of the sample frame have aided in reducing sampling errors and in the process, have reduced coverage errors.[9,11] Despite the efforts of the statistical agency to reduce sampling errors, content errors have still been found to be present in the data; this is more so a gendered phenomenon. Inconsistencies in health data collected from males showed that data collected from them are not accurate and cannot be relied on. This raises the question of the incentive for males to truthfully report on their health.

Yates[21] purported that people can have motives that retard them from giving or concealing the truth. Studies

on reliability and validity of data sources in the developing nations continue to emphasize the reduction of coverage errors (ie, sampling errors). While this is important in data quality, content errors have been substantially left unexplored as a means of providing explanations for the low quality of data in those developing nations. In the Caribbean, like many developing countries, males are socialized to be strong, brave, macho, not show emotions, and not display weakness, which explains their unwillingness to visit health care institutions for mere checkups and openly speak about illness affecting them. The issue of a motive that would account for their unwillingness to speak the truth about health matters is, therefore, embedded in the culture and definition of illness and its interpretation about their status. Males are sufficiently socialized to suppress weaknesses, and within the context of those societies, they must exhibit to females that they are strong, brave, and healthy, which explains their incorrect response related to health matters when asked by females. Yates,[21] while not stating that those matters are synonymous with males, provided us with justification for low data quality in the event that those issues are present in a sample.

There are several reasons that may explain the problems with reliability and validity of the health data in the present study. There is the case of social desirability bias where the participants say whatever is required to get the approval of the interviewers. Some participants do not even care about getting social approval, they just want to help the researcher, so they tell the interviewers what they think will help the interviewers. The possibility of collecting inaccurate data increases when government agencies are involved because of the declining trust that citizens have in government and public institutions in Jamaica. The data given may reflect a rejection of governmental authority and status. The converse may also be true where a researcher who is unaffiliated to a government agency receives accurate data. There is also the issue of the time when the interviewer seeks to interview males because this can adversely affect the data provided if the interviewers are competing with the important social and recreational times and events of the men.

Males are culturally competitive, which makes for strength, dominance, physique, and endurance critical to composition[25] that will be used to indicate to other males that they are healthier, superior, and stronger than the next competitive male. The challenge, therefore, is how do researchers develop an approach to collect data from males in which they have no motive to conceal the truth, give accurate answers, and concurrently ensure that interviewers' biases can be eliminated, or minimized so that data quality is not reduced in the data collection process. The challenge of mailing questionnaires to males in developing countries, particularly in Jamaica, is that the response rate would be very low and possibly so minimal that data analysis becomes problematic in providing pertinent information. The low reliability and validity of health data collected from males pose much public health context, as they experience the greatest mortality and not understanding their health is to further challenge public health practitioners and policy makers in instituting measures that will mitigate against their well-being.

To our knowledge, a mail survey has never been conducted in Jamaica. We have reservations about the likely success of this kind of survey as mentioned before, especially with men who are already underreporting their health status. Despite our reservations, the best way to know if a mail survey will work is to do one. However, there should be in-built incentives to increase the response rate. The PIN number for a specified amount of cell phone credit (in dollars) should be sent to the cell phone of participants whose completed questionnaires are received in the mail. The foregoing possibility highlights the fact that telephone surveys are also an underutilized method of data collection in a country where cell phone usage is widespread. The use of the cell phone has the advantage of allowing the participants to talk to the interviewers at any place and time that is convenient to them, which should improve the response rate. The validity and reliability would also be enhanced if the telephone interviewer is male.

Survey researchers more often do not use a mixed method research design. Sometimes, the discrepancies within a survey and between surveys are reduced by qualitative methods, such as individual interviews, focus groups, and participant observation among other methods. These methods will repopulate and enrich the text by writing back the individuals and their characteristics into published research while maintaining the use of regular statistical procedures.[44] No research method has a monopoly on reality so researchers should be eclectic in their methodological approach while being mindful that a bundle of techniques is not synonymous with intellectual sophistication and clarity.[45]

Reliability and validity can also be enhanced and discrepancies can be explained by the recognition that keeping things simple is best and doing less is more; there is greater clarity in using fewer variables for more highly targeted research problems. Health researchers should also be willing to question what was taught about the existing research methodologies and statistics.[46] In addition, future health researchers should pay attention to the role of mediator and moderator variables

in influencing the relationship between the independent and dependent variables because the measured influence of mediator and moderator variables can sometimes explain the discrepancies between surveys.[47]

Conclusion

The current study finds that there are many inconsistencies in health status data collected from older men in Jamaica. Generally, while this work did not examine males in Jamaica, using statistics from the Ministry of Health (Jamaica), it appears that the findings can be extrapolated to males. The wider implications for the findings are the challenges of using self-evaluated health data from males in planning their health and that currently we do not understand men's health. In researching men's health, it is not simply validating the instrument, but there are challenges in data collection that are unresolved and they increase nonsampling errors. Public health practitioners use self-reported health data from the national surveys and other subnational surveys, and they should understand the challenges faced in interpreting health data on males. Quality health data from males cannot produced using self-reporting of their health status in national or subnational surveys, as a part of this problem is the data collectors. Studies have not examined the influence of sex composition on the inaccuracies in health data, and this is clearly causing some noise in health statistics. The quality of health data in Jamaica, and by extension all nations, is influenced by the attitudes of respondents toward data collectors, the circumstances surrounding the interviewer, and culture and belief of the respondents. These continue to interfere with the quality of health data, which are still understudied in the Caribbean, and pose problems in understanding men's health. This should be a public health concern like epidemics, pandemics, and sanitary because poor-quality data will affect policies, programs, and implementations of strategies and alleviate particular health concerns faced by people, particularly males. Improvement in the quality of life through better health care must also integrate better-quality data collection because quality care requires accurate health data.

Jamaica is a middle-income developing nation in the Caribbean that has been collecting data for centuries, and it has 20-year data on health status. The efficiency and accuracy of statistical techniques used in gathering data are important in data collection. The PIOJ and the STATIN have continued to modify their sampling frame in an effort to reduce sampling errors. The widening and updating of the sampling frame have reduced coverage errors, as more people will be captured in national sample surveys. The current study has found that still there are errors in the quality of the health data collected from males, despite updating the sampling frame in 2007 in an effort to attain completeness of data coverage. Despite the aforementioned errors, the quality of the national survey, within the context of this study, are moderately good, and care should be taken in interpreting health data for males owing to the inconsistencies, which emerged from this study.

It is clear from the inconsistencies in the health data collected by the relevant agencies that the reliability of self-reported health data from males will pose a problem in public health planning. Sample surveys are used for teaching health care professionals; examining health care staff requirements; community health care; planning health care; planning and determining future care of patients; evaluation of public health policies; health care interventions; construction of community centers, hospitals and public clinics; and clinical and health service provisions. Two other issues that emerged from the present findings are 1) dichotomizing self-evaluated health for males loses some of the original information and 2) a sample of St Catherine is not the same as sampling the nation, and so a sample of the parish of St Catherine does not reliably reflect the detailed characteristic of the wider Jamaican population. Thus, care should be taken in the use of subnational samples to generalize about a population, and more so when it comes to data collected from males with regard to their health. Clearly, there are inconsistencies in health data collected from men in surveys, and this needs to be factored into their health intervention, and planning for their health status. These findings can inform further surveys and should stimulate an approach of how to collect data from males on their health status. If public health is to rely on research in order to effectively implement and attain its objectives, the data collected should be reliable and valid, and the current findings must be taken into consideration in aiding the process.

Acknowledgment

The authors thank the Data Bank in Sir Arthur Lewis Institute of Social and Economic Studies, the University of the West Indies, Mona, Jamaica, for making the data set (JSLC, 2007) available for use in this study.

Disclosure

The authors report no conflict of interest in this work.

References

1. Cox PR. *Demography*. 5th ed. London, UK: Cambridge University Press; 1976.
2. Preston SH, Heuveline P, Guillot M. *Demography: Measuring and Modeling Population Processes*. Oxford, UK: Blackwell Publishers; 2001.

3. Bryan T. Basic sources of statistics. In: Siegel JS, Swanson DA, editors. *The Methods and Materials of Demography*. 2nd. London, UK: Elsevier; 2004:9–39.
4. Newell C. *Methods and Model in Demography*. New York, NY: The Guilford Press.
5. Spiegelman M. *Introduction to Demography*. 6th ed. Cambridge, MA: Harvard University Press.
6. Morgan O, editor. *Health Issues in the Caribbean*. Kingston, Jamaica: Ian Randle; 2005.
7. Fowler FJ. *Survey Research Methods*. Thousand Oaks, CA: Sage Press; 1993.
8. World Health Organization. Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, and 1946, June 19–22; signed on 1946, Jul 22, by the representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and entered into force on Apr 7, 1948. Constitution of the World Health Organization, 1948. In: *Basic Documents*. 15th ed. Geneva, Switzerland: WHO; 1948.
9. Planning Institute of Jamaica (PIOJ) and Statistical Institute of Jamaica (STATIN). *Jamaica Survey of Living Conditions, 1988–2007*. Kingston, Jamaica: PIOJ and STATIN; 1989–2008.
10. Statistical Institute of Jamaica. *Jamaica Survey of Living Conditions, 2007* [Computer file]. Kingston, Jamaica: Statistical Institute Of Jamaica [producer], 2007. Kingston, Jamaica: Planning Institute of Jamaica and Derek Gordon Databank, University of the West Indies [distributors], 2008.
11. Statistical Institute of Jamaica. *Demographic Statistics, 2003*. Kingston, Jamaica: Statistical Institute of Jamaica; 2004:120–125.
12. Wilks R. Hypertension in the Jamaican population. Presented at the Trinidad and Tobago National Consultation on Chronic Non-communicable Diseases; 2007.
13. Jackson M, Walker S, Forrester T, Cruickshank J, Wilks R. Social and dietary determinants of body mass index in Jamaican of African. *Eur J Clin Nutr*. 2003;57:621–627.
14. Kahneman D, Riis J. Living, and thinking about it, two perspectives. Quoted in: Huppert FA, Kaverne B, Baylis N. *The Science of Well-Being*. Oxford, UK: Oxford University Press; 2005.
15. Finnas F, Nyqvist F, Saarela J. Some methodological remarks on self-rated health. *Open Public Health J*. 2008;1:32–39.
16. Helasoja V, Lahelma E, Prattala R, Kasmel A, Klumbiene J, Pudule I. The sociodemographic patterning of health in Estonia, Latvia, Lituania and Finland. *Eur J Public Health.* 2006;16:8–20.
17. Molarius A, Berglund K, Eriksson C, et al. Socioeconomic conditions, lifestyle factors, and self-rated health among men and women in Sweden. *Eur J Public Health.* 2007;17:125–133.
18. Leinsalu M. Social variation in self-rated health in Estonia: a cross-sectional study. *Soc Sci Med*. 2002;55:847–861.
19. Idler EL, Benjamin Y. Self-rated health and mortality: a review of twenty-seven community studies. *J Health Soc Behav*. 1997;38:21–37.
20. Idler EL, Kasl SV. Self-ratings of health: do they also predict change in functional ability. *J Gerontol: Soc Sci.* 1995;50B:S344–S353.
21. Yates F. *Sampling Methods for Census and Surveys*. 2nd ed. New York: Macmillan; 1981.
22. Kish L. Sampling and censuses. *Int Stat Rev*. 1979;47:99–109.
23. Godambe VP. A unified theory of sampling from finite populations. *J R Stat Soc*. 1955;B17:269–278.
24. Brewer KRW, Hanif M. *Sampling with Unequal Probabilities*. New York, NY: Springer-Verlag; 1983.
25. Chevannes B. *Learning to Be a Man: Culture, Socialization and Gender Identity in Five Caribbean Communities*. Kingston, Jamaica: The University of the West Indies Press; 2001.
26. Bourne PA. Predictors of good health status of rural men in Jamaica. *Calicut Med J*. 2009;7:e2,1–20.
27. Low W-Y, Chirk-Jenn NG, Choo, W-Y, Hui-Meng T. How do men perceive erectile dysfunction and its treatment: A qualitative study on opinions of men? *Aging Male*. 2006;9:175–180.
28. Ali M, de Muynck A. Illness incidence and health seeking behaviour among street children in Rawalpindi and Islamabad, Pakistan – a qualitative study. *Child Care Health Dev*. 2005;31:525–532.
29. Jamaica, Ministry of Health. Annual Report 2004. Kingston, Jamaica: MoHJ; 2005.
30. Jamaica, Ministry of Health. Annual Report 2006, 2007. Kingston, Jamaica: MoHJ; 2006, 2008.
31. Mackenbach JP, van de Bos J, Joung IM, van de Mheen H, Stronks K. The determinants of excellent health: different from the determinants of ill-health. *Int J Epidemiol*. 1994;23:1273–1281.
32. Manderbacka K, Lahelma E, Martikainsen P. Examining the continuity of self-rated health. *Int J Epidemiol*. 1998;27:208–213.
33. Manor O, Matthews S, Power C. Dichotomous or categorical response: analysing self-reported health and lifetime social class. *Int J Epidemiol*. 2000;29:149–157.
34. Bourne PA. A theoretical framework of good health status of Jamaicans: using econometric analysis to model good health status over the life course. *North Am J Med Sci*. 2009;1(2):86–95.
35. Hutchinson G, Simeon DT, Bain BC, Wyatt GE, Tucker MB, LeFranc E. Social and health determinants of well-being and life satisfaction in Jamaica. *Int J Soc Psychiatry*. 2004;50(1):43–53.
36. Bourne PA. An epidemiological transition of health conditions, and health status of the old-old-to-oldest-old in Jamaica: a comparative analysis. *North Am J Med Sci*. 2009;1:211–219.
37. Bourne PA. Good health status of older and oldest elderly in Jamaica: are there differences between rural and urban areas? *Open Geriatr Med J*. 2009;2:18–27.
38. Bourne PA. Social determinants of self-evaluated good health status of rural men in Jamaica. *Rural Remote Health.* 2009;9:1280.
39. Bourne PA, Rhule J. Good health status of rural women in the reproductive ages. *Int J Collab Res Intern Med Public Health*. 2009;1(5):132–155.
40. Bourne PA, McGrowder DA. Rural health in Jamaica: examining and refining the predictive factors of good health status of rural residents. *Rural Remote Health*. 2009;9(2):1116.
41. Bourne PA. Dichotomizing poor self-reported health status: using secondary cross-sectional survey data for Jamaica. *North Am J Med Sci*. 2009;1(6):295–302.
42. Kish L. *Survey Sampling*. New York, NY: Wiley; 1995.
43. Wilks R, Younger N, Mullings J, et al. Factors affecting study efficiency and item non-response in health surveys in developing countries: the Jamaica national healthy lifestyle survey. *BMC Med Res Methodol*. 2007;7:1–14.
44. World Health Organization. *Improving Data Quality: A Guide for Developing Countries*. Geneva, Switzerland: WHO; 2003.
45. Billig M. Repopulating the depopulated pages of social psychology. *Theory Psychol*. 1994;4:307–335.
46. Tukey JW. We need both exploratory and confirmatory. *Am Stat*. 1980; 34:23–25.
47. Cohen J. Things I have learned so far. *Am Psychol*. 1990;45:1304–1312.