Supplementary Table 1: Reasons for exclusion from sampling frame

Reason for exclusion	Number of households
Vacant[1]	586
House demolished	8
Occupants overseas	58
Error in the household listing[2]	68
Occupant not resident in Barbados	31
Under construction	8
Deceased	10
No resident older than 24 years	3
Participant hospitalised	3
Participant left address[3]	19
Other[4]	64

[1]Field team verified with neighbours that the house is unoccupied.

[2]Household was included in the original listing, but was found to be an error

[3]New address could not be located in postal service database.

[4]For example: occupant in hospital, prison, incapacitated.

Appendix 1: Additional information on non-responders and weighting scheme

NON-RESPONDERS VS RESPONDERS

Mean age of non-responders was slightly lower than that of those who took part in the survey: 49.9 years for non-responders, compared with 51 years for survey participants. Survey participants were also more likely to be female than non-responders: 62% compared with 50%.

When possible, the HotN survey collected age and sex information for households that did not take part in the survey. However, it should be noted that this information was only available for 50% of non-responding households. Therefore, it may be more appropriate to compare the age and sex distribution of the sample to that of the Barbados population, as shown in supplementary table 2.

Supplementary table 2: Percentage undercount (less than 100%) or percentage overcount (more than 100%) of the Health of the Nation (HotN) risk factor survey compared to the 2010 Barbados population and housing census

Age group	Men % undercount/overcount	Women % undercount/overcount
25-29	30.8	52.4
30-34	44.5	103.3
35-39	68.6	104.9
40-44	87.4	134.2
45-49	85.4	117.4
50-54	100.9	116.2
55-59	91.9	152.8
60-64	95.3	147.4
65-69	160.4	93.6
70-74	87.2	139.1
75-79	121.1	125.8
80-84	113.9	161.8
85+	63.6	89

WEIGHTING SCHEME

Principles guiding the Health of the Nation survey weights

Weights for disproportionate sampling are relatively non-controversial, but weights to adjust for non-response biases are largely dependent upon judgment. The HotN survey team has adopted the following principles for weighting survey data. These are:

1. The Enumeration District (the Primary Sampling Unit or PSU of the survey) has been selected using probability proportional to size (PPS) by the Barbados Statistical Service. This means that the probability of selecting each chosen household in the sample is the same for all EDs.

2. Non-equal probabilities of selection (caused by choosing 1 participant per household) are dealt with by applying weights proportional to the inverse of the probability of selection.

3. Non-response is calculated by ED. It is adjusted for by calculating the proportion of known eligible participants that refused to take part. This participant or 'unit' non-response is not the same as item non-response, where a participating individual may choose not to provide information for selected survey questions.

4. For a number of households, eligibility remained unknown (for example, the house was known to be occupied but the occupants could not be contacted). The percentage eligibility among these households was assumed to be the same as for households of known eligibility. This allowed an 'unknown eligibility' adjustment to be applied to all weights.

5. After selection and non-response adjustments have been made a final 'post-stratification' adjustment is applied. This final adjustment standardizes the age/sex distribution in the HotN sample to the known Barbados population.

Health of the Nation National Survey Weights: Details

Sampling weights in the HotN survey account for sample design features (such as selection of participants with unequal probabilities) and attempt to account for imperfections in the HotN sample (such as participant non-response) that might lead systematic departures of the HotN sample estimates from true population values. In other words, the purposes of weighting are:

- To compensate for unequal probabilities of selection.
- To compensate for (unit) non-response.
- To adjust the weighted sample distribution for key variables of interest (in the case of HotN, age and sex) to force it to conform to a known population distribution.

i. *Base-weight*

A base-weight for each sampled unit corrects for unequal probabilities of selection. The base-weight of each sampled unit is the reciprocal of its probability of selection into the sample. If a unit is included in the survey sample with probability P_i, then its base weight, denoted by w_i, is given by:

$$p(ij) = p(i) * p(j)$$

Appendix 1: Additional information on non-responders and weighting scheme

$$w(ij) = 1 / p(ij)$$

In the HOTN survey, EDs were selected with probability proportional to size, and households selected at an appropriate rate within EDs, such that every selected household had an equal probability of being selected. This means that ED and household base-weights are all the same, and the sample is referred to as self-weighting. However, the survey then chooses one participant per household, which nullifies the self-weighting design. So, the base-weight consists of a single adjustment to account for a 'participant weight' due to a single participant selected randomly (using Kish Tables) from several eligible participants in the household.

ii. *Nonresponse weights*

It is rarely the case that all desired information is obtained from all participants in a survey. It is important to keep survey nonresponse as low as possible, in order to reduce the possibility that the survey estimates could be biased in some way by failing to include (or including a disproportionately small percentage of) a particular portion of the population. The procedure of adjusting sample weights for nonresponse is the preferred practice for coping with unit nonresponse in major household surveys throughout the world. Essentially, the adjustment transfers the base-weights of all eligible non-responding sampled units to the responding units, and is implemented in the HOTN survey using the following steps:

Step 1: Apply the initial base-weights

Step 2: Partition the sample into 45 ED subgroups and compute weighted response rates for each ED

Step 3: Use the reciprocal of the subgroup response rates for non-response adjustments

Step 4: Calculate the non-response adjusted weight for the i-th unit as: w(i) = w1i * w2i, where w(1i) is the base-weight and w(2i) is the nonresponse adjustment weight.

Adjusting sample weights for unknown eligibility

During the HOTN survey, the eligibility of some households has been in question. In such cases, it is not known whether the dwelling unit is occupied or not. If it is actually occupied, then it should be classified as a non-responding dwelling unit. Otherwise, it is ineligible to be counted as a sample unit. Sometimes, survey teams assume that if no one is found in a dwelling unit during repeated visits, then that dwelling unit is unoccupied and hence ineligible. This is, in general, an incorrect assumption; one that often leads to erroneously inflated response rates. In the HotN survey we have been careful not to make this assumption, and have instead employed the following weight adjustment to account for households of unknown eligibility.

When the eligibility of some sampled dwelling units is unknown, their weights have been adjusted to account for this fact. The idea is to make some assumptions that permit the estimation of the

proportion of dwelling units with unknown eligibility that are actually eligible. We have used a simple approach, taking the proportion of sampled dwelling units known to be either eligible or ineligible, and applying that proportion to those of unknown eligibility. In the HotN survey, there are 2,076 households with the following response dispositions:

HotN households approached and response disposition

	Response	# households
1.	Complete interviews	1,234
2.	Eligible non-respondents	358
3.	Ineligibles	83
4.	Unknown eligibility	401
5.	TOTAL	2,076

The proportion of dwelling units of known eligibility that are actually eligible is (234+358)/(1234+358+83) = 0.9504.

We can therefore assume that the same proportion (0.95) of the dwelling units with unknown eligibility can be considered eligible. In other words, 95% of the 401 dwelling units with unknown eligibility (or 381 dwelling units) are actually eligible. We then adjust the weights of the eligible dwelling units (completed interviews and eligible nonrespondents) using an adjustment factor defined as follows:

F(ue) = sum(c)w(b) + sum(nr)w(b) + E*sum(ue)w(b) / sum(c)w(b)

> + sum(nr)w(b)

where E denotes the proportion of the unknown eligibility cases that are estimated to be eligible (E =0.95 for HotN). The summations over c, nr, and ue in the above formula denote, respectively, the sum of the base weights of dwellings with complete interviews, with eligible non-respondents, and of unknown eligibility. The adjusted base weights of dwellings with complete interviews and eligible non-respondents are then obtained by multiplying their initial base weights w(b) by the factor F(ue).

Non-coverage weights (also known as post-stratification weights)

Non-coverage refers to the failure of the sampling frame to cover all of the target population, leaving some sampling units with a reduced probability of selection into the sample selected for the survey. This is just one of many possible deficiencies of sampling frames used to select samples for surveys in developing countries. The HOTN uses a highly clustered ED design (approximately 2,000 households approached from just 45 EDs) to achieve its sample, and although this brings cost savings, the high clustering may lead to a non-representative sample.

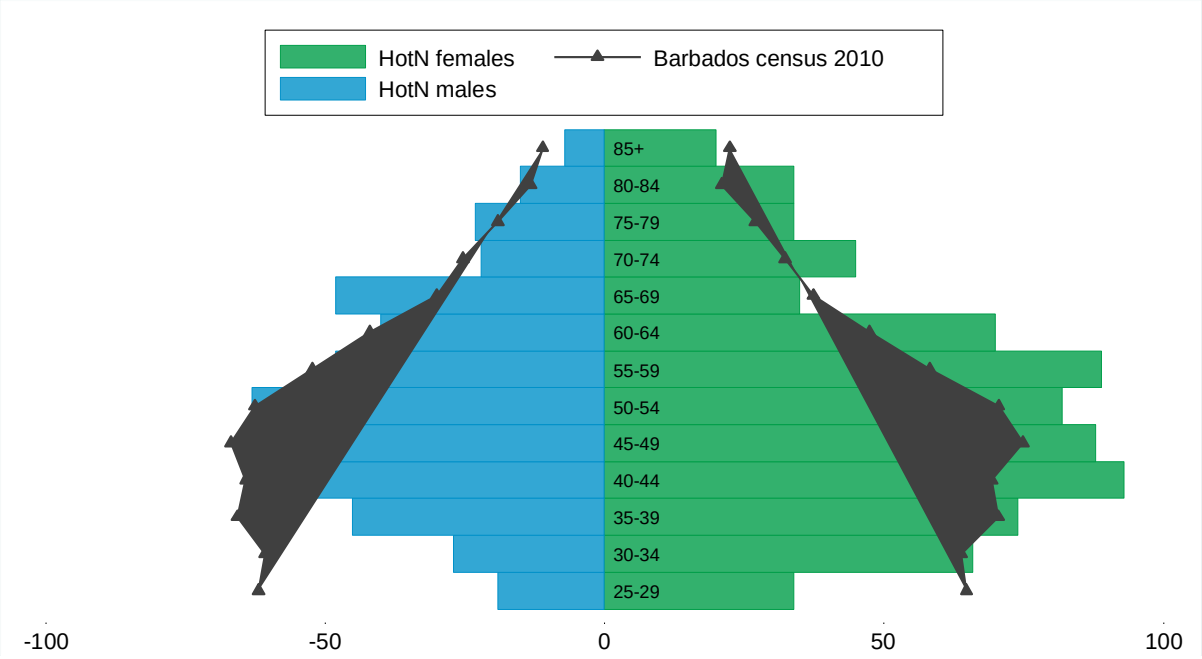
During a survey, there are field procedure considerations for minimizing non-coverage, but here we shall concentrate on post-hoc compensation for non-coverage through a statistical adjustment of the weights. In Barbados, reliable control totals are available for the entire population and for specified subgroups of the population (using the Barbados census), so we can adjust the weights of the sample units in such a way as to make the sum of weights match the control totals within the

specified subgroups. The subgroups are called post-strata, and the statistical adjustment procedure is called post-stratification. The post-stratification adjustment adjusts the weighted sampling distribution for certain variables so as to conform to a known population distribution.

In a sense, this procedure simultaneously compensates for non-response and non-coverage. However, this adjsutment is only within age/sex subgroups, whereas the non-response adjustment is within finer ED stratifications, may be prefereable, but is certainly different. To allow independence of nonresponse compensation and non-coverage compensation we have calculated and applied an item non-response adjustment separately to the various post-stratification adjustments.

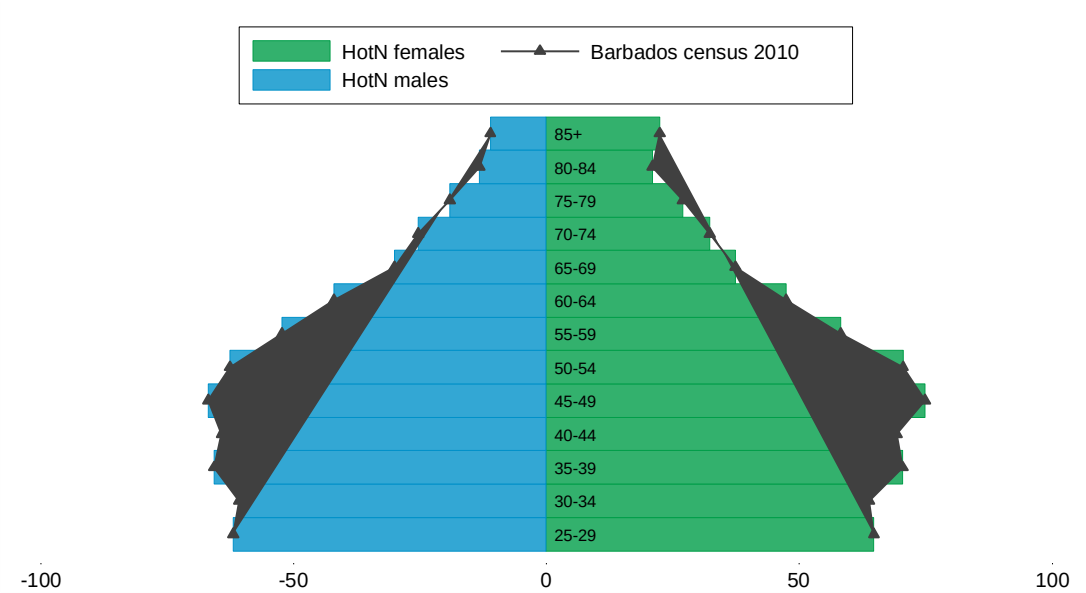
Survey age and sex participant distribution versus Barbados population distributions

The population pyramid figure generated below highlights the non-coverage (by age and sex) of the HotN survey sample compared to the Barbados population estimates from the 2010 census. It highlights the importance of accounting for under-representation of the young and over-representation of the elderly using post-stratification weighting by age and sex.



Appendix 1: Additional information on non-responders and weighting scheme

The population pyramid figure generated below shows the revised age/sex distribution of the HotN survey participants after post-stratification weights have been applied to follow the age/sex distribution from the Barbados Census 2010.



Weights applied in the paper: The social distribution of diabetes, hypertension, and related risk factors in Barbados: a cross-sectional study

Table 1: Selected sociodemographic characteristics of the survey sample, by sex and collectively. A single weight variable was created by multiplying base and non-response weights to adjust for participant selection probability and non-response together.

For all prevalence and prevalence ratio estimates, a combined single weight was applied, which accounts for base and non-response weights, and applying a post-stratification weight to standardize to the age/sex distribution of the Barbados Census 2010.

Supplementary table 3: Missing data in women – figures are n (%) for each category

	Diabetes	Hypertension	Tobacco use	Heavy episodic alcohol consumption	Low F&V intake	Physical Inactivity	Obesity	Raised cholesterol
Education								
< Secondary	9 (5.1)	0	0	0	8 (5.3)	1 (0.7)	9 (6.0)	16 (10.7)
Secondary	25 (8.5)	0	0	1 (0.3)	3 (1.0)	1 (0.3)	23 (7.8)	29 (9.8)
Higher technical	6 (6.3)	0	0	0	1 (0.8)	1 (0.8)	9 (7.3)	7 (5.7)
University	23 (11.8)	0	0	1 (0.5)	0	1 (0.5)	12 (6.2)	22 (11.3)
Occupation								
Manual	13 (10.6)	0	0	0	5 (4.1)	0	9 (7.3)	17 (13.8)
Intermediate	21 (8.2)	0	0	0	1 (0.4)	1 (0.4)	18 (7.0)	21 (8.2)
Professional	8 (5.6)	0	0	1 (0.7)	0	2 (1.4)	11 (7.6)	9 (6.3)
Not employed	12 (7.9)	0	0	1 (0.7)	4 (2.6)	1 (0.7)	11 (7.2)	13 (8.6)
Other	9 (10.1)	0	0	0	2 (2.3)	0	4 (4.5)	14 (15.7)
All women	63 (8.3)	0	0	2 (0.3)	12 (1.5)	4 (0.5)	53 (6.9)	74 (9.7)

Supplementary table 4: Missing data in men – figures are n (%) for each category

	Diabetes	Hypertension	Tobacco use	Heavy episodic alcohol consumption	Low F&V intake	Physical Inactivity	Obesity	Raised cholesterol
Education								
< Secondary	7 (6.6)	0	0	2 (2.4)	9 (11.0)	1 (1.2)	3 (3.7)	5 (6.1)
Secondary	25 (11.9)	0	0	4 (1.9)	3 (1.4)	2 (1.0)	8 (3.8)	27 (12.8)
Higher technical	4 (7.1)	0	0	3 (3.8)	4 (5.0)	0	3 (3.8)	7 (8.8)
University	5 (5.2)	0	0	0	1 (1.0)	1 (1.0)	1 (1.0)	6 (6.2)
Occupation								
Manual	16 (8.7)	0	0	3 (1.6)	7 (3.4)	1 (0.5)	7 (3.8)	17 (9.2)
Intermediate	7 (7.5)	0	0	1 (1.1)	4 (4.3)	2 (2.2)	3 (3.2)	8 (8.6)
Professional	4 (5.1)	0	0	1 (1.3)	1 (1.3)	1 (1.3)	1 (1.3)	4 (5.1)
Not employed	10 (20.4)	0	0	3 (6.1)	5 (10.2)	0	2 (4.1)	11 (22.5)
Other	4 (6.2)	0	0	1 (1.5)	0	0	2 (3.1)	5 (7.7)
All men	41 (8.7)	0	0	9 (1.9)	17 (3.6)	4 (0.9)	15 (3.2)	45 (9.6)