Available online at www.sciencedirect.com

Public Health

journal homepage: www.elsevier.com/puhe

Short Communication

Intracluster correlation estimates from a World Health Organisation STEPwise approach to surveillance (STEPS) survey for cardiovascular risk factors in Vellore, Tamil Nadu, India

A.M. Oommen [a,*], G.K. Mini [b,c], K. George [a]

[a] Community Health Department, Christian Medical College, Vellore, Tamil Nadu, India
[b] Sree Chitra Tirunal Institute for Medical Sciences and Technology, Trivandrum, Kerala, India
[c] Global Institute for Public Health, Ananthapuri Hospitals and Research Institute, Trivandrum, Kerala, India

ARTICLE INFO

Article history:
Received 16 July 2018
Received in revised form
13 November 2018
Accepted 21 December 2018
Available online 6 February 2019

Keywords:
Intracluster correlation
Surveys
Sampling

ABSTRACT

Objectives: Most World Health Organisation (WHO) STEPS surveys use cluster sampling to assess the prevalence of risk factors for non-communicable diseases (NCDs) for which design effects need to be estimated using intracluster correlation (ICCs) coefficients, for sample size calculation. Although there are many reports of risk factor surveys reported from developing countries, there are very few reports of ICCs for risk factors for NCDs, which can inform planning the appropriate sample size needed for such surveys. This study reports the ICCs for NCD risk factors, obtained from a WHO STEPS survey conducted in Vellore district, in the state of Tamil Nadu, South India.
Study design: Cross-sectional study.
Methods: A cross-sectional study was carried out in 48 urban clusters (wards) and nine rural clusters (villages) between 2011 and 2012, using the WHO STEPS methodology for assessing behavioural, anthropometric, physical and biochemical risk factors. The ICC estimates for various risk factors were obtained using loneway and xtmelogit commands using STATA to study clustering of risk factors.
Results: The number of respondents was 6196 adults aged 30–64 years. The median ICC of cardiovascular risk factors in the urban area was 0.046, while it was 0.064 in the rural area. Clustering was higher for behavioural risk factors such as physical activity (ICC: 0.179 rural, 0.049 urban) and fruit and vegetable intake (ICC: 0.105 rural, 0.091 urban) as compared with physical risk factors (ICCs for hypertension: 0.044 rural, 0.006 urban; body mass index: 0.046 rural, 0.041 urban) and biochemical outcomes such as fasting plasma glucose (ICC: 0.017 rural, 0.027 urban).

* *Corresponding author*. Community Health Department, Christian Medical College, Vellore, Tamil Nadu, 632002 India. Tel.: +91 9486368796.
E-mail address: anuoommen@cmcvellore.ac.in (A.M. Oommen).
https://doi.org/10.1016/j.puhe.2018.12.019

Conclusions: This study provides estimates of ICCs for cardiovascular risk factors from Vellore, South India, as such data have not been reported from WHO STEPS surveys in India or neighbouring countries. Such estimates of ICCs if reported from various WHO STEPS being carried out across the country can contribute to better planning of epidemiological surveys. Clustering of behavioural risk factors at village/ward level as seen in this study points to the need for community-based interventions for health promotion, as spatial clustering influences behaviour, which in turn affects chronic disease outcomes.

Introduction

Many studies have estimated prevalence of cardiovascular risk factors using World Health Organisation (WHO) STEPS methodology.[1–4] The WHO recommends selection of samples using multistage cluster sampling methodology for STEPS surveys.[1] In cluster surveys, increasing the sample size obtained based on simple random sampling by the design effect or variance inflation factor is to be considered to compensate the loss of variability that occurs as a result of cluster sampling.[1,5] It is important to adjust for clustering in both the design and analysis stages to avoid overestimation of effective sample size by using statistics meant for simple random sampling. However, few WHO STEPs studies have reported intracluster correlation (ICC) estimates in the analysis, especially from developing countries such as India, with the possibility of overestimation of the effective sample size.

Factors influencing variability of individuals within clusters depend on cluster definition and characteristics (e.g. availability of fruits and vegetables, access to avenues for engaging in physical activity and availability of healthcare services). Estimates of ICCs for various cardiovascular risk factors are not available from India to calculate design effects for cluster-based surveys. The objective of this analysis was to estimate ICCs from a cardiovascular risk factor survey carried out in Vellore district, Tamil Nadu, South India (2011–2012), which aimed to study the prevalence of coronary heart disease and its risk factors.[4]

Methods

A cross-sectional WHO STEPS survey was carried out among adults aged 30–64 years to assess prevalence of cardiovascular risk factors in urban and rural Vellore, Tamil Nadu, for which the results related to prevalence have been previously published.[4] The current article describes the ICCs obtained from this study as additional information to enable planning of cluster surveys for non-communicable disease (NCD) risk factors.

Ethical clearance was obtained from the Institutional Review Board and Ethics Committee, and written informed consent was obtained from participants.

The rural component of the study was carried out in a rural block where the predominant occupations are non-agricultural and agricultural manual labour and in 48 urban wards of the Vellore Corporation in 2011–2012, a tier-II city. Nine villages (clusters) were selected in the block based on simple random sampling, from the list of 23 villages in the block where a similar study had been carried out in 1991–1994.[4] All adults aged 30–64 years were included in the study, constituting nine rural clusters of varying sizes (mean size: 420: range 63–798). In urban Vellore, from the list of streets in each ward, one street (cluster) was randomly selected in each of the 48 wards. In each selected street, all adults aged 30–64 years in the first 40 consecutive households were included for the study. As the number of participating adults among the 40 households in the selected streets varied, these urban clusters were unequal (mean size: 50, range 16–72).

Sociodemographic and behavioural risk factors were gathered through home visits using the STEPS interview schedule, while physical and biochemical factors (STEPS 2 and 3) were collected at special early morning study clinics, after an overnight fasting of at least 8 h.[4] Hypertension was defined as elevated blood pressure measurements ($\geq$140 mm Hg systolic or $\geq$ 90 mm Hg diastolic) or on medication and diabetes as fasting plasma glucose (FPG) of $\geq$126 mg/dl or on medication.[1] Abdominal obesity was defined as waist circumference of $\geq$80 cm in women and $\geq$90 cm in men.[4] Current tobacco use was taken as current smoking or use of smokeless tobacco, and current alcohol use was defined as use of an alcoholic drink within the last 12 months. A serving of fruits/vegetables was taken as equal to 80 g.[1]

ICCs with 95% confidence intervals were calculated for urban and rural areas separately, considering villages and urban wards as clusters, using the 'loneway' (large analysis of variance, random effects) and xtmelogit followed by estat icc commands, using STATA v. 13.1 (Stata Corporation, College Station, Texas, USA) for continuous and binary variables, respectively.[6,7]

Results

The number of subjects interviewed for behavioural risk factors was 6196 (rural 3799), which represented 83% and 77% of the population aged 30–64 years in the study area.[4] The response of participants for biochemical tests was lower, with the sample for these tests representing 68% and 60% of the eligible population in the rural and urban areas, respectively. The mean age in the nine rural clusters ranged from 44.4 years to 46.9 years, while in the 48 urban clusters it ranged from 42.2

years to 47.4 years. The average cluster size was 422 in the rural area and 50 in the urban areas. Females formed 56% of the sample in both urban and rural areas.

The ICCs for various behavioural, physical and biochemical variables are shown in Table 1. The median of all ICCs for the urban data was 0.046 (median for continuous variables was 0.048 and for binary variables 0.045). The median value of all the rural ICCs was 0.064 (median for continuous variables was 0.064 and for binary variables 0.079), showing a higher level of clustering of cardiovascular risk factors in villages as compared with urban wards.

The ICCs for behavioural risk factors were high, especially in the rural area (ICC for physical activity: 0.179; fruits and vegetable intake: 0.105) (Table 1). The ICCs for hypertension (rural 0.044, urban 0.006), body mass index (rural 0.046, urban 0.041) and FPG (rural 0.017, urban 0.027) were lower compared with behavioural risk factors (Table 1). The ICC values were low (<0.1) for most clinical outcome variables, with abdominal obesity being an exception (Table 1).

Discussion

This study documents ICCs for NCD risk factors, from a large cross-sectional WHO STEPS survey from Vellore, South India,[4] according to the guidelines suggested by Campbell et al.[11] namely description of the data set, method of calculation of ICCs and precision of ICCs. Although it is recommended that researchers publish the ICCs obtained in their cluster-based studies to enable better estimates of ICCs for various locations,[12] this is rare in literature from developing countries.[13]

In this study, the highest ICCs were observed in the rural area for behavioural indicators such as daily physical activity and servings of fruits and vegetables. A similar result was seen from a study in the United States which showed high clustering of unhealthy diet within physician practice areas.[8] As these behavioural factors are linked to issues of accessibility and availability, such as access to fruits and vegetables or availability of opportunities for physical activity, they appear to be clustered, in rural areas. This is also an important point to bear in mind while designing intervention programs for diet and physical activity for which community-level interventions need to be planned rather than only individual-based programs. The high clustering of alcohol use, especially in urban areas (ICC 0.055), as shown in this study implies that this behaviour is also influenced by within cluster characteristics such as availability and social norms. This was also reported from a population-based study in Madhya Pradesh, where the ICC for alcohol use scores was 0.052, with a wide variation of alcohol use scores between wards/villages.[14]

Table 1 – Intracluster correlation (ICC) coefficients for cardiovascular risk factors.

Risk factor	Rural (clusters = villages)			Urban (clusters = wards)		
	Number	ICC	95% confidence interval	Number	ICC	95% confidence interval
Behavioural						
Physical activity minutes/day	3774	0.179	0.0–0.343	2390	0.049	0.022–0.077
Fruit and vegetable servings per day (1 serving = 80 g)	3782	0.105	0.0–0.211	2379	0.091	0.049–0.133
Current tobacco use	M: 1666	M: 0.028	M: 0.008–0.100	M: 1058	M: 0.044	M: 0.015–0.124
	F: 2129	F: 0.069	F: 0.016–0.249	F: 1339	F: 0.226	F: 0.016–0.418
Alcohol use in males[a] in the last 12 months	1666	0.019	0.006–0.063	1057	0.055	0.022–0.128
Physical						
Systolic blood pressure (mm Hg)	3317	0.038	0.0–0.081	1938	0.017	0.0–0.033
Diastolic blood pressure (mm Hg)	3317	0.034	0.0–0.073	1938	0.023	0.004–0.042
Percent with hypertension[b]	3317	0.044	0.015–0.124	1959	0.006	0.0–0.101
Body mass index (BMI) (kg/m^2)	3315	0.046	0.0–0.096	1938	0.041	0.015–0.067
Overweight[c]	3315	0.070	0.026–1.78	1938	0.046	0.024–0.089
Waist circumference (cm)	M: 1417	M: 0.090	M: 0.0–0.185	M: 805	M: 0.027	M: 0.0–0.062
	F: 1857	F: 0.174	F: 0.013–0.336	F: 1111	F: 0.115	F: 0.057–0.173
Percent with abdominal obesity[d]	M: 1417	M: 0.104	M: 0.036–0.261	M: 805	M: 0.043	M: 0.014–0.124
	F: 1857	F: 0.182	F: 0.076–0.376	F: 1111	F: 0.141	F: 0.071–0.259
Biochemical						
Fasting plasma glucose (FPG) (mg/dl)	3095	0.017	0.0–0.039	1857	0.027	0.006–0.048
Percent with diabetes[e]	3104	0.042	0.018–0.139	1879	0.023	0.006–0.077
Total cholesterol (mg/dl)	3095	0.099	0.0–0.200	1825	0.097	0.051–0.144
Triglycerides (mg/dl)	3095	0.012	0.0–0.29	1825	0.048	0.019–0.077
High-density lipoprotein (HDL) (mg/dl)	3095	0.064	0.0–0.133	1825	0.374	0.268–0.481

M: males; F: females; FPG: fasting plasma glucose.
[a] < 1% of females were current alcohol users.
[b] Blood pressure ≥140/90 mm Hg or on medication.
[c] Percent with BMI ≥25 kg/m^2.
[d] Waist circumference ≥80 cm (females) and ≥90 cm (males).
[e] FPG ≥126 mg/dl or on medication.

As individuals within a cluster (e.g. a village) are likely to interact with one another more than with other clusters leading to ICC,[5] interventions which affect lifestyle may also be spread within a cluster, leading to similar behaviour. The high clustering of hypertension in rural areas (ICC 0.046) compared with urban areas (ICC 0.006) was also seen in a study from Tanzania, which reported an ICC of 0.167 for hypertension in rural areas compared with 0.056 in urban areas.[13] Clustering for diabetes was low in this study from Vellore (ICC for FPG: 0.017 rural, 0.027 urban), as was also found in Tanzania (ICC for HbA1c: 0.012).[13]

The lower ICCs for clinical outcome variables as opposed to 'process variables' have been shown previously, and behavioural factors are possibly more clustered as they are processes which lead to physical and biochemical outcomes.[9,15] A study in Canada among 84 primary-care practices also found similar low ICCs of <0.1 for most clinical outcomes relating to cardiovascular disease and its risk factors.[15]

The design effect has not been presented here as it depends on cluster size and is not as generalisable as the ICC.[10] The sample population for this study was from a single district in South India and hence is generalisable only to places with similar sociocultural settings. However, as burden of disease and risk-factor patters in India are highly variable from state to state, it is useful to map local variations in disease and risk-factor patterns. The age group for the study was restricted to 30–64 years based on recommendation for such risk factor surveys, and therefore results may not be the same for all age groups.

Conclusion

Although the WHO suggests a blanket design effect of 1.5 for STEPS surveys,[1] it would also be useful for countries to have a 'catalogue of ICCs' as suggested by Littenberg and Maclean, based on estimates from large surveys.[6] Using local ICCs, where available, may improve calculation of design effect and sample size estimation for cluster-based studies compared with relying on external estimates of ICCs, leading to over or underestimation of sample size requirements.[15] The estimates of ICCs in this study could encourage the calculation of design effects for WHO STEPS surveys in similar settings in developing countries, enabling researchers publish more statistically authentic findings.

Author statements

Ethical approval

The study was approved by the Institutional Review Board and Ethics Committee of the institution which conducted the study, and written informed consent was obtained from participants.

Funding

The cross-sectional WHO STEPs survey was funded by the Indian Council of Medical Research (ICMR), which was involved in planning the survey, monitoring and interpreting major survey findings.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

A.M.O. was involved in conception and design of the study, survey implementation, analysis and interpretation of data and writing the paper. G.K.M. was involved in design of the study, analysis and interpretation of data and critically reviewing the draft paper. K.G. was involved in design of the study, interpretation of data and critically reviewing the draft paper. All authors approved the final version of the study.

REFERENCES

1. World Health Organization. The WHO STEPwise approach to noncommunicable disease risk factor surveillance. Available from: http://www.who.int/chp/steps/STEPS_Manual.pdf?ua=1 Accessed on June 20th, 2018.
2. Thankappan KR, Shah B, Mathur P, Sarma PS, Srinivas G, Mini GK, et al. Risk factor profile for chronic non-communicable diseases: results of a community-based study in Kerala, India. *Indian J Med Res* 2010;**131**:53–63.
3. Tripathy JP, Thakur JS, Jeet G, Chawla S, Jain S, Pal A, et al. Prevalence and risk factors of diabetes in a large community-based study in North India: results from a STEPS survey in Punjab, India. *Diabetol Metab Syndrome* 2017;**9**:8. https://doi.org/10.1186/s13098-017-0207-3.
4. Oommen AM, Abraham VJ, George K, Jose VJ. Prevalence of risk factors for non-communicable diseases in rural and urban Tamil Nadu. *Indian J Med Res* 2016;**144**:460–71.
5. Donner A, Klar N. *Design and analysis of cluster randomized trials in health research*. London: Arnold Publishing Co; 2000.
6. Littenberg B, MacLean CD. Intra-cluster correlation coefficients in adults with diabetes in primary care practices: the Vermont Diabetes Information System field survey. *BMC Med Res Methodol* 2006;**6**:20.
7. Eldridge SM, Ukoumunne OC, Carlin JB. The intra-cluster correlation coefficient in cluster randomized trials: a review of definitions. *Int Stat Rev* 2009;**77**:378–94. https://doi.org/10.1111/j.1751-5823.2009.00092.x.
8. Thompson DM, Fernald DH, Modl JW. Intraclass correlation coefficients typical of cluster-randomized studies: estimates from the robert wood johnson prescription for health projects. *Ann Fam Med* 2012;**10**:235–40.
9. Campbell MK, Fayers PM, Grimshaw JM. Determinants of the intracluster correlation coefficient in cluster randomized trials: the case of implementation research. *Clin Trials* 2005;**2**:99–107.
10. Gulliford MC, Ukoumunne OC, Chinn S. Components of variance and intraclass correlations for the design of community-based surveys and intervention studies: data from the Health Survey for England 1994. *Am J Epidemiol* 1999;**149**:876–83.
11. Campbell MK, Grimshaw JM, Elbourne DR. Intracluster correlation coefficients in cluster randomized trials: empirical insights into how they should be reported. *BMC Med Res Methodol* 2004;**4**:9.

12. Killip S, Mahfoud Z, Pearce K. What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. *Ann Fam Med* 2004;**2**(3):204–8.
13. Stanifer JW, Egger JR, Turner EL, Thielman N, Patel UD. Comprehensive kidney disease assessment for risk factors, epidemiology, knowledge, and attitudes (CKD AFRiKA) study. Neighborhood clustering of non-communicable diseases: results from a community-based study in Northern Tanzania. *BMC Public Health* 2016;**16**:226.
14. Rathod SD, Nadkarni A, Bhana A, Shidhaye R. Epidemiological features of alcohol use in rural India: a population-based cross-sectional study. *BMJ Open* 2015;**5**:e009802.
15. Singh J, Liddy C, Hogg W, Taljaard M. Intracluster correlation coefficients for sample size calculations related to cardiovascular disease prevention and management in primary care practices. *BMC Res Notes* 2015;**8**:89.