



Cluster randomisation and design effects in school surveys: an exploration of the Irish Health Behaviours in School-aged Children data.

Title	Cluster randomisation and design effects in school surveys: an exploration of the Irish Health Behaviours in School-aged Children data.
Author(s)	Molcho, Michal;Kelleher, Cecily C.;Nic Gabhainn, Saoirse;Walsh, Kieran
Publication Date	2008-09
Publisher	National Institute of Health Sciences

Population Health

title	Cluster Randomisation and Design Effects in School Surveys: An Exploration of the Irish Health Behaviour in School-Aged Children Data
author	Walsh, K., ¹ Molcho, M., ¹ Dineen, B., ² Kelleher, C., ³ NicGabhainn, S. ¹ Health Promotion Research Centre, NUI, Galway ¹ Centre for Health Services Research and Development, NUI Galway ² School of Public Health and Population Science, University College Dublin ³

INTRODUCTION

School surveys are typically conducted in classrooms where children are clustered and thus each participating student is not fully independent of others. Such unavoidable cluster randomisation increases the size of standard errors, resulting in widened confidence intervals, larger probability values and therefore reduced statistical power and effective sample size. The need to address these issues is becoming more evident in the literature, with increasing impetus to publish on the effects of such sampling techniques and to consider more fully the analysis implications for complex survey data. A first step is to assess the design effects (DEFF), which is the ratio of the number of participants required to achieve precision using cluster randomisation to the number that would be required if all participants were independent.

OBJECTIVE

These analyses investigated the presence and magnitude of design effects in the Irish Health Behaviour in School-aged Children (HBSC) study.

METHODOLOGY

The Irish HBSC sample was constructed by stratifying school lists by region (n=8), randomly selecting schools within regions and then randomly selecting a class from within each school year. Therefore, classrooms served as the cluster or primary sampling unit (PSU). The sampling design was based on a 95% confidence interval of $\pm 3\%$ around a proportion of 50% and an estimated design factor of 1.2. The data were collected through self-completion questionnaires administered by teachers during a class period.

The dataset consisted of 2,365 males and 3,411 females post-primary school students. Males ranged in age from 11.2 years to 19.3 years, with a mean of 15.1 years and a SD of 1.4. Females ranged in age from 12.0 to 20.3 years, with a mean of 15.1 years and a SD of 1.4. All cases with missing data on stratification level (regions) or primary sampling unit (class) were deleted. Separate data files were created for each item. Each file was imported into Epi Info 6. The complex sample function was then used to compute the DEFF for each item, with region and class identifiers set a 'Strata' and 'PSU' variables.



RESULTS

Although the analysis was completed for all 189 survey items, only the DEFF values of those variables that were featured in the last international HBSC report are presented in Table 1.

Table 1 - DEFF values of selected HBSC survey items

Survey item	DEFF
Perceived academic achievement	1.655
Self-rated health	1.486
Body image (perception of being fat)	1.056
Level of agreement with having a good life	1.044
Life satisfaction	1.031
Level of agreement that family is well-off	2.435
Occupational status of father	1.750
Level of agreement that the local area is well-off	1.556
Frequency of going to bed or school hungry	1.098
Attempting to lose weight	1.815
Frequency of eating vegetables	1.456
Frequency of drinking 'diet' soft drinks?	1.420
Frequency of drinking whole fat milk	1.273
Frequency of eating fruit	1.190
Frequency of snacking at weekend	1.187
Frequency of eating sweets	1.006
Age of onset of menstruation (girls only)	10.986
Frequency of moderate to vigorous physical activity	1.383
Frequency of drunkenness	6.069
Frequency of drinking alcohol	5.590
Frequency of cannabis use	4.338
Frequency of smoking tobacco	3.548
Frequency of being bullied	1.869
Frequency of being injured	1.681
Teachers treating students fairly	2.911
Students enjoying being together	2.532
Liking school	2.350
Teachers encouraging students to express views	1.796
Students accepting me as I am	1.750
Frequency of e-communication with friends	3.280
Ease of talking to mother	2.135
Frequency of evenings spent with friends	1.529

CONCLUSIONS

DEFF values vary substantially across survey items. Variables that measure school context, friendship groups and risk behaviours display the largest DEFF values. These analysis serve as insight into appropriate sampling requirements for subsequent HBSC and other school based surveys.

REFERENCES

Available on request.

FUNDING

HBSC Ireland is funded by the Health Promotion Policy Unit and the Office of the Minister for Children, Department of Health and Children. These specific analyses were funded as part of a research project grant from the Health Research Board.

