Catalogue no. 11-522-XIE

Statistics Canada International Symposium Series - Proceedings

Symposium 2006 : Methodological Issues in Measuring Population Health

2006

Statistics Canada   Statistique Canada

Canada

Proceedings of Statistics Canada Symposium 2006
Methodological Issues in Measuring Population Health

Estimating Sample Size for Complex Surveys: Building Consensus in an Environment of Multiple Hypotheses, Multiple Stakeholders and Limited Budgets

Victor JC, O'Connor S, Diemert LM, Bondy S, Brown KS, Cohen J, Ferrence R, Garcia J, McDonald P, Selby P, Stephens T[1]

Abstract

How does one efficiently estimate sample size while building concensus among multiple investigators for multi-purpose projects? We present a template using common spreadsheet software to provide estimates of power, precision, and financial costs under varying sampling scenarios, as used in development of the Ontario Tobacco Survey. In addition to cost estimates, complex sample size formulae were nested within a spreadsheet to determine power and precision, incorporating user-defined design effects and loss-to-followup. Common spreadsheet software can be used in conjunction with complex formulae to enhance knowledge exchange between the methodologists and stakeholders; in effect demystifying the 'sample size black box'.

KEY WORDS: Sample size, Population-based Survey, Study Design.

1. Introduction

Few researchers would disagree that adequate sample size is critical to good study design. Sample size calculations for experimental and observational studies are often easily performed through use of standard formulae and sample size software. However, deriving appropriate sample size estimates for a study is complicated when the study is examining multiple outcomes and/or multiple exposures. The utility of standard formulae and software is limited when determining sample size for broad surveys of the general population. These surveys may be designed for multiple purposes such as developing population estimates of both outcomes and exposures, and testing hypotheses. Often, these surveys also involve multiple investigators, with different research foci.

We present the approach taken by the Ontario Tobacco Research Unit (OTRU) to develop the design and sample size required for its Ontario Tobacco Survey (OTS). The OTS is a regionally-stratified telephone survey of over 7500 adult smokers and non-smokers residing in Ontario. The survey incorporates a rolling longitudinal component that surveys smokers, and a cross-sectional component surveying non-smokers. We describe this approach as a potential guide for other research groups undertaking the development of a large-scale, multipurpose survey.

2. Background

The Ontario Tobacco Research Unit is an academic unit with offices at the Centre for Addiction and Mental Health, the University of Toronto and the University of Waterloo in Ontario, Canada. The research direction of the unit is managed by a team of seven principal investigators.

---

[1]J. Charles Victor, Shawn O'Connor, Lori M. Diemert, Susan Bondy, K. Stephen Brown, Joanna Cohen, Roberta Ferrence, John Garcia, Paul McDonald, Peter Selby, Thomas Stephens, Ontario Tobacco Research Unit, 33 Russell St. Toronto, Ontario, Canada M5S 2S1. Corresponding author's email: charles_victor@camh.net.

The Ontario Tobacco Research Unit is tasked with providing timely research on the status of tobacco use in the province of Ontario. In addition to providing annual estimates of tobacco-related outcomes and exposures, OTRU also evaluates the policies and initiatives of the Government of Ontario to stem tobacco use in the province. To accomplish these goals, OTRU has, for years, used provincial surveys (e.g., Centre for Addiction and Mental Health Monitor (CAMH Monitor) and Ontario Student Drug Use Survey) and national health surveys (e.g., Canadian Community Health Survey and Canadian Tobacco Use Monitoring Survey). However, these surveys often did not fully meet OTRU's needs. The Ontario based surveys are focussed on a range of mental health issues and addictions, providing only moderate content for tobacco-related questions. National surveys either provided small sample sizes for Ontario, or did not provide adequate tobacco-related content. As the prevalence of smoking declines in Ontario (Ontario Tobacco Survey Research Unit, 2006) the data from these surveys are increasingly insufficient to perform sub-group analysis. As a result, OTRU committed to developing an Ontario-based population survey that would both provide periodic estimates of tobacco-related outcomes and exposures, and enable longitudinal hypothesis testing with adequately large sample sizes to allow small sub-group analyses.

3. Planning the Design and Determining Sample Size

The first step in the development of the OTS was to determine the design. Several options were considered including fully cross-sectional, fully longitudinal, and a combined longitudinal/cross-sectional hybrid. Given the multiple goals of the survey and the different research foci of the seven principal investigators of the research unit, it was decided to combine a longitudinal component with a cross-sectional component. However, the precise nature of the two components, and stratification of the sampling frame remained difficult to define. The barriers to moving forward with the design of the survey included determining the power to evaluate multiple hypotheses while remaining within our budget allocation. Several sample size calculations proved unfruitful in the design process, as these calculations were static, and did not address the questions of all investigators.

3.1 Development of a Design and Sample Size Tool

Due to the complex longitudinal/cross-sectional hybrid design, we began work on a spreadsheet to estimate costs for given designs and numbers of participants (Figure 1: all figures at end of text).

This first screen of the tool allowed the user to input the number of smokers sampled into a longitudinal component and additionally into a cross-sectional component (if required), as well as the number of non-smokers sampled into a cross-sectional component. The tool was designed to account for user-defined loss-to-follow-up rates, and whether losses would be replaced in subsequent waves. Finally, the 'Design' page allowed the user to enter the estimated cost per interview, and thus the cost over the study period.

One goal of the survey was to determine statistically stable population estimates of tobacco-related outcomes and exposures in specific age groups and regions. A decision was made to sample within regions of Ontario, defined by telephone area codes, in order to achieve adequate sample in the northern regions of the province. Furthermore, the decision was made not to sample based on age, but to choose a sample size that would provide sufficient numbers of participants between the ages of 18 and 30 years. To estimate the number of participants in each smoking status by age by region subgroup, a second page was added to the design tool that estimated the allocation of the sample size in each of these strata based on 2001 census information and smoking prevalence from the 2002 CAMH-M (Figure 2). While not visible in Figure 2, several stratification scenarios were evaluated on this page (e.g., combining 416 and 905 telephone exchanges vs. separating these exchanges). The estimates provided on the second page of the tool are linked with the user-defined inputs from the first page allowing real-time evaluation of different sample sizes and their impact on the age, regional, and smoking status make up of the final sample.

To evaluate the ability of the cross-sectional components of the survey to provide precise and stable population estimates of tobacco-related outcomes and exposures, a third page was added to the tool (Figure 3). Formulae to calculate the width of 95% confidence intervals for estimates of proportions of 5%, 25% and 50% were integrated into the page. Due to the complex nature of the sampling for the survey, traditional asymptotic confidence interval formulae would provide biased and often smaller widths. Thus the formulae used were adjusted according to Lohr (1999) to account for the estimated design effects (DEFFs). An input field on the page allowed the user to enter the estimated DEFFs and view changes to the precision of estimates. Moderate values between 1.5 and 2.0 were used for estimating precision for the OTS survey (Lohr, 1999). This page stratified estimates by smoking status and age group, allowing users to see the precision of analyses for sub-groups of particular interest.

While the third page of the tool ("Precision), was integral in evaluating the ability of the survey to provide estimates, a fourth and fifth page were added to calculate power for research hypotheses. The fourth page calculated power for both user-defined and preset cross-sectional odds ratios incorporating the user-defined design effects linked from the previous page (Figure 4). The final page was used to calculate power for longitudinal analyses (two timepoints), for both binary and continuous outcomes (Figure 5); again, these formulae incorporated the user-defined DEFFs entered on the third page (Lohr, 1999; Rosner, 1995).

4. Discussion

Developing a complex survey reporting multiple population estimates and testing multiple hypotheses while being cost-effective requires a multi-faceted approach to sample size estimation. However, the development of a sample size tool such as that developed for the OTS is not always warranted. Indeed commercial sample size calculators and/or good references for sample size forumlae (see Fleiss, 2003; Lohr, 1999; Rosner, 1995) may be all that investigators need during the initial planning phase of a project. However, these common tools fail for many reasons on more complex research studies such as general population surveys. First, these tools make it difficult to readily examine several hypotheses at the same time. Secondly, the interface often precludes simple real-time adjustments to parameters and revised results. That is, the sample size programs, and certainly statistical formulae are not easily accessed nor understood by all investigators. Often there needs to be a facilitator (the analyst), to guide the derivation and interpretation of results. Finally, there is frequently a disconnect between the investigators that are concerned about the budgetary constraints of the research project, and the analyst who is concerned with developing an accurate estimate of sample size required to achieve the greatest statistical power. By incorporating sample size and power formulae into common spreadsheet software along with design information and assumptions (e.g., loss-to-follow-up), and cost information, one is able to marry all of these aspects into one tool that is accessible to most investigators.

We found the most effective use of this tool was at team meetings where, equipped with a laptop computer and projector, the analyst was able to use the tool and alter the parameters of design and sample size according to multiple team members' questions. The provision of immediate responses expedited the design and sample size process. Possibly the most novel and important feature of the tool was the integration of the attrition and cost estimate function. These functions made it possible to immediately evaluate the financial feasibility of design suggestions and requested for sample sizes, while ensuring that adequate power was maintained. Following team meetings, the sample size tool was made available to all team members. Given its common spreadsheet platform, this made the tool usable by all investigators that wished to examine different design scenarios outside of the group meetings.

5. Recommendations

For research teams considering large projects that may have multiple hypotheses or purposes, multiple investigators, and limited budgets we strongly suggest the development of a tool similar to that which is presented here. We advise the development of a common spreadsheet tool where the analyst incorporates:

1) The possible designs of the study
2) User-defined estimates of attrition and cost
3) Published sample size formulae (include design effects where appropriate)

We advise the use of this tool in team meetings, where the real-time capabilities of this tool can have the greatest impact. This particular tool is not designed to calculate sample size calculations for highly complex analyses (e.g. multi-level modelling). The greatest utility to these tools is most likely for broad surveys of the general population. Nonetheless, many complex sample size formulae could easily be incorporated into a tool such as this with the proper expertise.

References

Fleiss, JL., Levin B., Paik MC. (2003), *Statistical Methods for Rates and Proportions 3rd ed.*, New York: Wiley.

Lohr S. (1999), *Sampling: Design and Analysis*, Pacific Grove CA: Duxbury Press.

Ontario Tobacco Research Unit (2006). *Indicators of Smoke-Free Ontario Progress*. [Special Reports: Monitoring and Evaluation Series, 2005-2006 (Vol. 12, No. 2)]. Toronto, ON: Ontario Tobacco Research Unit.

Rosner B. (1995), *Fundamentals of Biostatistics*, Boston: Duxbury.

Figure 1: Design and Cost Calculations of Spreadsheet Tool to Develop Sample Size for Complex Projects

	A	B	C	D	E	F	G	H	I
1	ENTER VALUES HERE	Total Longitudinal Smoking Sample	XS Smoking Sample Wanted per Wave	XS Non-Smoking Sample Wanted per Wave	Loss to Follow-up Rate/mth	Replace loss-to-follow-up?	Approx. Cost / Interview		
2		1000	1000	1000	30	Y	40		
3		Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6		
4	Longitudinal core - SMOKERS								
5	Initial Call	500	500	0	0	0	0		
6	Replacements for sample 1		150	150	150	0	0		
7	Replacements for sample 2			150	150	150	0	Longitudinal	
8	Baseline interview	500	650	300	300	150	0	Cohorts	
9	Followup1		350	455	210	210	105	Cohort 1	
10	Followup2			245	319	147	147	Cohort 2	
11	Followup3				172	223	103	Cohort 3	
12	Followup4					120	156	Cohort 4	
13	Subtotal (LS Smokers)	500	1000	1000	1001	850	511	Cohort 5	
14									
15	Cross-sectional SMOKERS								
16	Obtained from LS	500	650	300	300	150	0		
17	Additional calls	500	350	700	700	850	1000		
18	Subtotal (XS smokers)	1000	1000	1000	1000	1000	1000		
19									
20	Cross-sectional NON-SMOKERS								
21		1000	1000	1000	1000	1000	1000		
22	Subtotal (Non-smokers)	1000	1000	1000	1000	1000	1000	# Interviews	14,962
23								Total Cost	$598,480.00
24	Total interviews in time period	2000	2350	2700	2701	2700	2511	Cost per 12 mth	$199,493.33
25	Total interviews in Fiscal Year	2000	5050		5401		2511		
26	Projected Cost (by Fiscal Year)	$80,000.00	$202,000.00		$216,040.00		$100,440.00		
27									
28									

Page1 - Design / Page 2 - Allocation / Page 3 - CIs / Page 4 - XS_Power & ORs / Page 5 - Long Power

Figure 2: Allocation of Sample Size Across Statification Scenarios

	A	B	C	D	E	F	G	H	I
1	Using 5 Area Code Regions								
2		REGION		807/705				613	
3				Age				Age	
4	Smoking Status		18-29	30+	Total		18-29	30+	Total
5	Smoker	Daily	84	91	175		84	89	173
6		Occasional	16	9	25		16	11	27
7	Non-Smoker	Former	11	45	56		10	37	47
8		Never	89	55	144		90	63	153
9	Total		200	200	400		200	200	400
10									
11		REGION		416/647				905/289	
12				Age				Age	
13	Smoking Status		18-29	30+	Total		18-29	30+	Total
14	Smoker	Daily	73	85	158		79	88	167
15		Occasional	27	15	42		21	12	33
16	Non-Smoker	Former	11	25	36		9	33	42
17		Never	89	75	164		91	67	158
18	Total		200	200	400		200	200	400
19									
20		REGION		519				ALL REGIONS	
21				Age				Age	
22	Smoking Status		18-29	30+	Total		18-29	30+	Total
23	Smoker	Daily	84	90	174		404	443	847
24		Occasional	16	10	26		96	57	153
25	Non-Smoker	Former	10	39	49		51	179	230
26		Never	90	61	151		449	321	770
27	Total		200	200	400		1000	1000	2000

28 *Occasional Smokers Defined using OTRU definition of those smoking > 1 and < 30 cigarettes in the past 30 days + > 100 cigarettes in lifetim

29 **Calculations of Region Specific allocations based on smoking status proportions from CCHS 2001

Page1 - Design / Page 2 - Allocation / Page 3 - CIs / Page

Figure 3. Precision Estimates (95% Confidence Interval Widths) Based on Sample Size, Allocation and Design Effects

Enter DEFF			Precision for percentages. Width of 95% confidence bands around percent		
2.0					
			Expected Values		
			5%	25%	50%
Provincial Estimates			±	±	±
Smoker		1000	1.9	3.8	4.4
	Daily Smoker	847	2.1	4.1	4.8
	Occasional Smoker	153	4.9	9.7	11.2
Non-Smoker		1000	1.9	3.8	4.4
	Former	230	4.0	7.9	9.1
	Never	770	2.2	4.3	5.0
Analyses Restricted to <30 yrs old			±	±	±
Smoker		500	2.7	5.4	6.2
	Daily Smoker	404	3.0	6.0	6.9
	Occasional Smoker	96	6.2	12.3	14.1
Non-Smoker		500	2.7	5.4	6.2
	Former	51	8.5	16.8	19.4
	Never	449	2.9	5.7	6.5
Analyses Restricted to 30+ yrs old			±	±	±
Smoker		500	2.7	5.4	6.2
	Daily Smoker	443	2.9	5.7	6.6
	Occasional Smoker	57	8.0	15.9	18.4
Non-Smoker		500	2.7	5.4	6.2
	Former	179	4.5	9.0	10.4
	Never	321	3.4	6.7	7.7

Page1 - Design | Page 2 - Allocation | **Page 3 - CIs** | Page

Figure 4: Power Calculation for Cross-sectional Study Odds Ratios Based on Sample Size, Allocation, and Design Effects

Power Calculations for simple OR's from XS study

Enter: Proportion of 'unexposed' having factor of interest in %	Enter: Odds Ratio	p(exp) %	Recall:	
5.0	2	9.5	XS Smokers	1000
			XS NonSmoker	1000

		n	Power	Examples of Minimum OR's (N=1000 Smokers and 1000 Non-Smokers)** Expected Outcome Rates: 5%	10%	25%
Provincial Estimates*						
Smoker		1000	78	2.02	1.71	1.48
	Daily Smoker	847	71	2.14	1.78	1.53
	Occasional Smoker	153	19	>4.00	3.43	2.62
Non-Smoker		1000	78	2.02	1.71	1.48
	Former	230	26	3.70	2.80	2.20
	Never	770	67	2.20	1.83	1.56
ENTIRE SAMPLE		2000	97	1.68	1.48	1.32
Analyses Restricted to <30 yrs old*		n	Power			
Smoker		500	49	2.60	2.08	1.73
	Daily Smoker	404	41	2.82	2.24	1.83
	Occasional Smoker	96	13	>4.00	>4.00	3.30
Non-Smoker		500	49	2.60	2.08	1.73
	Former	51	9	>4.00	>4.00	>4.00
	Never	449	45	2.70	2.15	1.78
ENTIRE SAMPLE		1000	78	2.02	1.71	1.48
Analyses Restricted to 30+ yrs old*		n	Power			
Smoker		500	49	2.6	2.08	1.73
	Daily Smoker	443	45	2.72	2.16	1.79
	Occasional Smoker	57	10	>4.00	>4.00	>4.00
Non-Smoker		500	49	2.60	2.08	1.73
	Former	179	21	>4.00	3.18	2.50
	Never	321	34	3.15	2.43	1.96
ENTIRE SAMPLE		1000	78	2.02	1.71	1.48

* All analyses are two-tailed, with alpha=0.05

** All analyses are two-tailed, with alpha=0.05 beta=0.80

Page 3 - CIs | **Page 4 - XS_Power & ORs** | Page 5 - Lc

Figure 5. Power Calculation for Longitudinal Differences in Binary and Continuous Outcomes Based on Sample Size, Allocation, and Design Effects

Analyses of Proportions			Analyses of Mean Change	
Proportion maintaining some factor b/w timepoints: (ie Proportion that respond yes or no to quit attempt at both T1 and T2) in %	Odds Ratio of Favourable Direction of Change (ie Odds of quitting smoking b/w T1,T2 compared to odds of taking up smoking among former smokers)		Mean Difference	Variance of Difference
85	2		7	139

Provincial Estimates*		n	Power to detect a difference: Power for OR	Power for Mean Difference
Smokers		1000	92	100
	Daily Smokers	820	91	100
	Occasional Smokers	180	81	100
Analyses Restricted to <30 yrs old*		**n**		
Smoker		500	87	100
	Daily Smoker	410	86	100
	Occasional Smoker	90	77	98
Analyses Restricted to 30+ yrs old*		**n**		
Smoker		500	87	100
	Daily Smoker	410	86	100
	Occasional Smoker	90	77	98

* All analyses are two-tailed, with alpha=0.05

Page 4 - XS_Power & ORs | Page 5 - Long Power