Improving the Sample Size Calculation Process for Peel Health: A Rapid Review of the Evidence

Sheila Datta, Epidemiologist
Karen Funnell, Health Analyst
Nancy Ramuscak, Supervisor

October 2012

Region of Peel
Working for you

Table of Contents

Key Take Home Messages

1. Before undertaking sample size calculations, public health teams must have a clear understanding of the research question to be examined. This will help determine whether a population-based approach or a power analysis approach should be used for sample size determination.
2. The Epidemiology team will change from providing “one-time” consultations with teams on sample size determination to an iterative process that will explore important aspects of the research project. Consultations will aim to improve the methodological design and quality of the proposed research activities with careful consideration of costs, time and resources available.
3. Given that estimates and sampling error cannot be computed for non-probability samples, the epidemiology team will not provide sample size calculations for surveys in which a sampling frame does not exist (or where non-probability sampling methods have been used).
4. Survey results using non-probability sampling should avoid making inferences and generalizations to the population of interest. Instead, results should be restricted only to those participating in the survey.
5. The Epidemiology team will consider strategies to improve the selection of the final sample (e.g. stratification, clusters) to enhance sample efficiency.
6. Estimations related to the population parameter and the survey design effect should be informed by the literature or derived from pilot surveys (where possible).

Executive Summary

Issue and Context

Sample size calculations are based on an understanding of the research objectives as well as theoretical and statistical assumptions. Errors or incorrect assumptions can lead to sample size estimates that are either too small or too large. An inadequate sample size leads to a lack of statistical power to detect significant differences between groups of interest while an overestimate of the sample size can lead to wasted public health resources. Epidemiologists are aware of the magnitude of project decisions which rest on sample size estimations – decisions related to the feasibility of undertaking the project, the available budget, and staff resources. Epidemiologists are largely afraid of making an error that will inadvertently set teams on an incorrect path for their project.

Research Question: *What specific elements are needed when planning sample size calculations related to the following three scenarios:*

1. *Population health surveys where multiple outcomes are examined;*
2. *Population health surveys where multiple outcomes are examined and where the sample selection involves clusters (e.g., schools)*
3. *Consultation requests that involve program planning and evaluation.*

Literature Search

A text word search of books was conducted using a number of library catalogues using the subject phrase “sample size”. The large number of books retrieved was deemed impractical for retrieving relevant evidence. A second search was conducted using previously acquired books and requesting recommendations from experts in the field. This search strategy yielded nine

results, including eight books and one chapter in a technical manual. Applying inclusion criteria specified for this rapid review yielded six texts that were critically appraised.

Critical Appraisal and Findings

Of the six textbooks reviewed, four were rated as “strong”, two “moderate” and none “weak”. Textbooks examined sample size calculations using either population-based approaches or approaches related to power analysis. Other fundamental elements affecting sample size estimates were discussed including reducing bias introduced from non-sampling errors, estimating the variability of an outcome of interest and determining the sample design effect.

Applicability and Transferability

Members agreed there is strong support to improve sample size calculations conducted by Peel Public Health. It was recognized that high quality research must be seen as ethical, justifiable and a good use of available resources. Endorsement and support from middle management was deemed important to engaging teams in the new process.

Conclusion

Findings of this rapid review have resulted in both methodological and organizational process changes related to sample size determination. Calculations will use either population-based or power analysis approaches depending on the goal of the project. A consultative process will allow teams to carefully consider research objectives as well as be cognizant of available budget, time and resources. The sample size consultation process will also support departmental efforts to build a culture of numeracy.

1 Issue

Sample size calculations for quantitative research projects require a careful balance between precision and operational constraints such as budget, resources and timelines.[1] Epidemiologists are assigned the task of determining the sample size required for research projects and surveys conducted by Peel Public Health staff. These calculations are conducted based on an understanding of the objectives and methods of the project, as well as theoretical and statistical assumptions related to the study design and sampling methodology. Errors or incorrect assumptions lead to sample size estimates that are either too large or too small. An inadequate sample size leads to imprecise estimates and a lack of power to detect significant differences between groups. An overestimate of the required sample size leads to a waste of resources to answer the research question. Given that the team conducts five to ten sample size calculations per year, this rapid review was required to address an important knowledge gap and to increase our confidence in sample size determination. It will also outline the information needed from program staff for accurate calculations.

Building a common understanding within the Epidemiology Team of the elements involved in sample size calculations will lead to an improved process for developing sample size calculations. The primary purpose of this report is to explore the following practice question: *What specific elements are needed when planning sample size calculations related to the following three scenarios:*

1. Population health surveys where multiple outcomes are examined

2. Population health surveys where multiple outcomes are examined and where the sample selection involves clusters (e.g., schools, workplaces, physician offices)

3. Consultation requests that involve program planning and evaluation

A secondary purpose is to inform the development of a standardized and documented process for correctly conducting this activity.

2 Anecdote

A request for a sample size calculation for a new Peel Public Health project causes stress and worry for the Epidemiologist responsible for completing it. The complexity of some of the projects and the magnitude of the decisions based on the result are often the main drivers behind this concern.

In 2009, Epidemiology was asked to determine the sample size required for municipal-specific estimates of breastfeeding at six months postpartum. After consultation, an estimated sample size of 975 mothers (367 from Mississauga, 364 from Brampton and 244 from Caledon) was provided to Family Health staff and recruitment for the survey began. This sample size was based on the assumption of probability sampling from an unspecified sampling frame, an estimated prevalence of the main outcome of interest of 50% and a response rate of 70%.

A number of the assumptions of the calculations were not met through the study design and recruitment methods. For example, mothers were not being randomly selected from a sampling frame but were rather a convenience sample of mothers contacted by telephone by the Healthy Babies Healthy Children assessment team. A previous Peel breastfeeding survey found that the

prevalence of the main outcome of interest (i.e., exclusive breastfeeding at six months) was approximately 20% but this knowledge was not used to reduce the sample size estimate. Finally, to recruit potential participants, consecutive mothers were asked to participate, without consideration for meeting the municipal-specific sample size estimates, which led to the recruitment of too many respondents from Brampton and Mississauga and not enough from Caledon.

As a direct result, municipal-specific estimates were not possible and therefore, the survey included many more respondents than would have been required for a Peel-level estimate only (790 participants vs. 375 participants). In addition, given the non-probability sampling method used, the representativeness of the sample to the overall Peel population was questionable. Family Health must use the findings of the Breastfeeding Duration Survey with caution until the implementation of a more rigorously designed survey.

3 Context

The Epidemiology Team provides consultation and technical support for population health research and program evaluations. Each Epidemiologist is assigned to support the work within a Division or teams within a Division. Teams expect the Epidemiologist to have a high level of knowledge about research methods and statistical analysis.

Epidemiologists are highly aware of the magnitude of the project decisions which rest on their sample size estimate – decisions related to feasibility of the project, budget, staff resources, and

timing of the project. Epidemiologists are largely afraid of making a mistake and inadvertently setting the teams on an incorrect path for their project.

For a number of reasons, the Epidemiologist does not always feel confident that the sample size estimate provided to the team is accurate. Sample size calculations are complex even for simple study designs and teams are increasingly posing more sophisticated research questions. Epidemiologists may not be aware or confident with the methods required for sample size determination for these complex projects. The tools available for sample size calculations for complex designs are not well understood. Epidemiologists are not always intimately involved in the project and are unsure of the validity of the assumptions they must make. Finally, the Epidemiologist may be unaware of changes to the data collection methodology during implementation. In some situations this will mean that the assumptions of the sample size calculation have been violated, making the sample size estimate no longer valid. Without a revised sample size calculation, the teams may collect data from too many or too few participants. The Epidemiologist may not be aware of this problem until the end of the project when it is too late to address it.

Epidemiologists typically work in isolation from their colleagues when completing sample size calculations. Another Epidemiologist, if asked to complete the same request, could make different assumptions which may potentially lead to a different sample size estimate.
If the Epidemiology Team provides an inaccurate sample size estimate, the result may be a lack of confidence in the abilities of the team to support Peel Public Health projects.

The Rapid Review research team conducted a conceptual model exercise with Internal Client Services (ICS), Communicable Disease Epidemiologists, and the Epidemiology team to set the parameters for this review and identify themes and issues relevant to the practice question (Appendix A1).

The researchers then developed an analytical map to thematically group key elements relevant to the practice question and to illustrate how different study designs can impact sample size calculations (Appendix A2).

4 Literature Search

Search Strategy

Summary level evidence was available for this practice question. Two methods were used to retrieve relevant evidence (Appendix B). First, the researchers and librarian conducted a text word search of books using a number of library catalogues using the subject phrase “sample size” and restricting findings to books published in English (Appendix B1). Due to the large number of books retrieved, this search strategy was deemed impractical for retrieving relevant evidence.

The researchers then conducted a second search which included reviewing previously acquired books and requesting recommendations from experts in the public health field affiliated with Public Health Ontario, the Dalla Lana School of Public Health and McMaster University.

The inclusion criteria used were: English language, the text covered at least one of the three scenarios included in the research question and the text provided sufficient detail regarding the process to be followed for sample size calculations within the scenario(s) covered. The exclusion criteria used were: books where more recent editions were available and books where sample size determination was not the main topic of interest.

The search strategy yielded nine results, including eight books and one chapter in a technical manual. Of the nine texts, three books did not meet the inclusion criteria specified for this review. Specifically, one book was excluded as a more recent edition was available and two books were excluded because they did not provide sufficient detail related to the practice question. The final retrieval process yielded six texts that were critically appraised.

Critical Appraisal

At the time of this review, no tool could be found to critically appraise books. The researchers adapted, with permission, criteria used by librarians at Cornell University as the basis for a critical appraisal tool (Appendix C).

The researchers reviewed and critically appraised the six texts. Each text was given an overall appraisal of “strong evidence”, “moderate evidence” or “weak evidence” and justification for the appraisal rating was documented. Of the six texts reviewed, four were rated as “strong”, two “moderate” and none “weak”. Therefore all six texts were included in this rapid review.

Description of Included Books

The six texts included for this Rapid Review can be divided into two approaches: 1) population-based approach and 2) power analysis. See "Synthesis of Findings" section for a more detailed description of these two approaches. A more detailed review of these texts can be found in Appendix D "Data Extraction Table".

Texts Using a Population-Based Approach

1) **Statistics Canada, *Survey Methods and Practices* (2003): Strong**

This manual, originally developed for Statistics Canada methodologists, provides a practical guide for survey planning, design and implementation. Chapter 8 discusses sample size calculations and provides insight into methods related to allocation and stratification. The book outlines a four-step practical approach to sample size calculations which can be applied to various scenarios presented in practice. The approach provides the necessary structure and process to guide practitioners' decisions and assumptions when undertaking such work.

2) **Levy and Lemeshow, *Sampling of Populations: Methods and Applications* (1999): Strong**

This book is a relevant introductory text to sampling principles and methods. The chapters are arranged by type of sampling (e.g., simple random sample, one-stage cluster, two-stage cluster) and within each chapter there is a short section dedicated to sample size determination. The text outlines a number of preliminary planning steps that should be conducted prior to survey implementation, including; develop objectives; decide on the population to be extrapolated; identify other potential methods for obtaining the information; identify the decisions that are to

be made with the results; and determine the required precision of the results, subgroups of interest and the available resources and time for the survey.

3) **Lemeshow S., Hosmer D.W., Klar J. & Lwanga S.K., *Adequacy of Sample Size in Health Studies* (1990): Strong**

This book was developed to address a number of typical questions posed by health practitioners concerning sample size determination and includes a number of illustrative examples. The book is structured into two parts: the first section emphasizes the "practical application" to address sample size requirements for various surveys and study designs; the second section discusses the underlying "theoretical process" of sample size determination. This text includes estimation for longitudinal, cross-sectional and case-control studies, as well as studies involving cluster sampling. The book uses simple language and is written for the practitioner with no previous experience in sampling principles and methods.

Texts Using Power Analysis Approach

4) **Cohen J, *Statistical Power Analysis for the Behavioural Sciences* (1988): Moderate**

This comprehensive examination of power analysis is often used as a foundation for other texts. The text describes power as a function of sample size, effect size and alpha (significance criterion). This text focuses on the underlying principles of power analysis and is organized by the type of statistical test being conducted (e.g., comparison of means, proportions, correlation). Each chapter includes a description of the statistical test of interest and the formula required to estimate the effect size, and has power tables and sample size tables for different levels of significance.

5) **Dattalo P, *Determining Sample Size: Balancing Power, Precision and Practicality*. (2008): Moderate-Strong**

This book presents three approaches which are research question-oriented: Power analysis, Confidence Intervals and Computer-Intensive Strategies. Based on the type of statistical test to be performed, the text presents formulae and examples for these approaches. However, in order to begin, a desired sample size based on feasibility or other studies must be proposed; once calculations are made, this sample size gets adjusted according to preferred levels of alpha or power. The book provides good descriptions of basic terms and concepts, along with suggestions and links to software and tools.

6) **Kraemer C. & Thiemann S., *How Many Subjects? Statistical Power Analysis in Research* (1987): Strong**

This book introduces a simple technique that allows researchers to compute approximate sample sizes and study power. Because the same technique can be used with only slight modifications for different statistical tests, researchers can easily compare the sample sizes required for different study designs and make cost-effective decisions when planning a study. The book addresses a gap in the applied statistics literature by emphasizing important principles of design, measurement and analysis that should be considered when determining the sample size needed for a study.

5 Synthesis of Findings

> *"You are making a fundamental flaw. You think that statistics has the answer. It is a balance between real-world challenges and statistical precision."*
> Dr. David Streiner.
>
> *"Determining sample size can be described as a patchwork quilt of procedures. No single software package exists that allows researchers to determine sample size according to these [true] strategies across all statistical procedures..."* [2]

Sample size calculations are done:

- to obtain a sample that is representative of the population of interest; or
- to ensure adequate power to detect differences between two or more groups of interest.

The practitioner must have a clear understanding of the purpose of the project to determine the most appropriate approach.

Of the six texts used for this rapid review, two described population-based approaches to selecting samples, [1, 3] three outlined approaches related to power analysis,[2, 5, 6] and one described both approaches.[4]

Population-based Approach to Sample Size Calculations

The population-based approach can include probability or non-probability sampling techniques. None of the texts used in this review discussed sample size determination for surveys involving non-probability sampling. Reliable estimates or estimates of sampling error cannot be calculated

for non-probability samples since elements of the population have an unknown chance of being selected.[1-3] Generalizing from a non-random sample to a population can yield biased conclusions.[2] Respondents selected for the sample could differ systematically from the population of interest.[2]

In probability sampling, every element in the population has a known chance of being selected for the sample.[2] Probability sampling techniques (such as simple random sampling) are used to ensure that the sample selected is reflective of the population to which one wants to generalize one's findings. Methodological techniques that can be used to improve on simple random sampling include stratification and cluster sampling. Stratification techniques may be used to ensure representation of sub-groups of interest.[2] Cluster sampling techniques can effectively reduce survey costs and increase sampling efficiency by reducing the resources required during the data collection phase.[2]

A researcher draws a sample from available records (the sampling frame) based on the sampling technique chosen (e.g., simple random sampling, cluster sampling). Statistics Canada recommends the use of a 4-step approach:

1. Conduct initial sample size calculation
2. Adjust for the size of the population
3. Estimate the design effect
4. Adjust for non-response rate.

In situations where there are multiple outcomes of interest, the sample size calculations are repeated for each outcome. The largest value obtained through this process determines the sample size required.

Power Analysis Approach

For public health research questions which explore differences between various groups, the power analysis approach should be used.[2, 5, 6] The particular formulae for power analysis will differ depending on the type of statistical test used to answer the research question. Once the type of statistical test has been determined, the effect size, significance criterion and required sample size can be used to ensure sufficient power to detect differences that may exist between groups. In general, when alpha is set to be 0.05, 80% power should be the goal of a research study.[2, 5] Decisions are guided by factors related to ensuring the sample size for specific sub-groups of interest are sufficiently large to detect statistical differences *a priori*.

The Dattalo text outlines the process to determine sample size requirements based on power analysis. A summary of this process is as follows:

1. Determine the research question to be answered and the type of statistical test to be conducted
2. Based on previous research studies or perceived feasibility, determine a reasonable sample size
3. Calculate the power to detect a difference between the groups of interest based on the type of statistical test required
4. Compare the power of the study to the desired level of power (80% typically)

5. Adjust the sample size either up or down to obtain adequate power.

Where there are multiple outcomes of interest, or more than one research question to be answered, the above process would be conducted for each of the outcomes/questions and the maximum sample size required would be used.

Please see Table 1 (p. 26) for a comparison of the population-based approach and the power analysis approach.

Fundamental Elements Affecting Sample Size Estimates

There are additional elements that were addressed in all texts reviewed.

Systematic Sampling Errors versus Nonsampling Errors

Sample size determination minimizes sampling errors that occur randomly.[1] However, it is unable to control for nonsampling errors (i.e., bias) related to the data collection methodology. Bias cannot be controlled for by increased sample size; strong research methods are required to minimize bias.[1]

Variability of the Outcome of Interest

The magnitude of the variability between subjects will affect the sample size necessary to achieve a given level of precision.[1] A large sample is required to accurately measure characteristics that have high degrees of variability. The required sample size is largest when the variability of the characteristics of interest is at its maximum (e.g., this occurs at 50% for a dichotomous variable). Statistics Canada recommends “that a conservative estimate of the

population variability of a characteristic be used when calculating the required sample size".[1] Thus, when there is no advanced knowledge, the outcome of interest should be set at 50% (for a dichotomous variable). Given that sample surveys measure more than one characteristic, each with differing variability, the sample size should be determined for the characteristic that is thought to have the highest variability, or for the one(s) deemed to be most important.[1]

Size of the Population

The size of the population has a large impact on sample size for small populations, a moderate impact for medium-sized populations, and a minor role for larger populations. For example, a population size of 50 would require 44 respondents to complete the survey for an outcome in which the true population proportion is 0.50. Conversely, the required sample size quickly approaches a maximum size of 400 for surveys drawn from populations larger than 5,000.[1] For a more detailed discussion please refer to the Statistics Canada manual, p.155-156.[1]

Design Effect

Cluster and stratified samples typically have sampling errors that are larger than simple random samples (SRS) of the same size.[2] This is because clusters or strata are typically homogeneous because of selective grouping effects. For example, students in a school tend to be similar in attitudes and behaviours or are exposed to common environments or influences compared to students outside of the school. Clusters and strata therefore require an adjustment to the variance and standard errors of the parameter estimate in order to account for their complex design.[1] This is achieved by calculating the "design effect", which is the ratio of the sampling variance of an

estimator under a given design to the sampling variance of an estimator under SRS of the same sample size.[1]

When a more complex sample design is used, the sample size required to satisfy a given level of precision is multiplied by the design effect.[1] For example, a design effect of 3 means that the sample size needs to be three times as large as it would be using SRS. In general, the larger the degree of homogeneity of people within clusters, the larger the design effect.[2]

The value of the design effect is equal to 1 for simple random sampling designs, less than 1 for stratified sample designs, and greater than 1 for cluster sampling designs.[1] In other words, stratified sampling reduces the sample size requirements while cluster sampling increases the sample compared to the sample size that would be required if using simple random sampling procedures.

An estimate of the design effect can be obtained from examining previous research that surveyed a similar population group on a similar topic or by calculating the design effect from a pilot survey.[1] It is more challenging to determine the appropriate design effect to use when there is no prior knowledge of the clustering effect on the sampling variance. In such cases, the design effect must be set to at least 2 but design effects of highly clustered designs may be as high as 6 or 7.

6 Applicability and Transferability

The Peel Public Health knowledge broker facilitated an Applicability and Transferability meeting to assess the feasibility (applicability) and generalizability (transferability) of the results

of this Rapid Review. In attendance were Megan Ward (AMOH), Julie Stratton (Manager, Epidemiology), Andrea James (Epidemiologist), JoAnne Fernandes (Epidemiologist), Lori Greco (knowledge broker) and the Rapid Review team. A summary of the points of discussion can be found in Appendix E.

The A&T meeting attendees agreed that there is strong support for the implementation of the Rapid Review findings, namely developing a process to improve sample size calculations and support for research conducted by Peel Public Health teams. The new process for sample size calculations and, more generally, increased direct support from Epidemiology for team research projects was seen as part of larger organizational changes that are underway which are placing more onus on teams to effectively plan new research projects (e.g., EIDM strategic priority, proposed Research Review Committee). High quality research must be seen as ethical, justifiable and a good use of available resources (i.e., accountable for public funds). Improving the quality of research and evaluations conducted by Peel Public Health will improve the profile of Epidemiology within Peel Public Health, but will also increase the profile of Peel Public Health within the Regional corporation and externally (e.g., other health units, community partners).

Epidemiology must develop a process which recognizes the time and financial pressures for teams, as well as takes into account the Epidemiology Team resources available for this work. At the beginning of each year, Epidemiology will work proactively with teams to identify priority research questions and will work with them to develop a strong research proposal, which can be implemented quickly in the event sufficient funding becomes available.

Endorsement and support for the new process from Public Health Management Team is required, although Epidemiology will work directly with middle management (supervisors and managers) to engage teams in the new process. The new process must be shown to have value to the teams and the benefits must be clearly identified in advance to engage teams in the process (i.e., avoid process being seen as “one more thing to do”).

The new process for working with teams should be pilot tested with a small number of teams, with a period of reflection afterwards on the lessons learned and identification of potential process improvements. Epidemiology must be cautious of “overselling” their services, meaning that they must carefully assess the resources available to support teams and ensure that they are able to meet other demands that fall within the epidemiologist scope of work (i.e., writing of health status reports, health status analysis).

Epidemiology will identify sources of external support for sample size calculations which are more complex; these sources may include no-cost support (e.g., biostatistician from PHO, researchers whose funding requires the provision of support to community agencies) and fee-for-service consultations (e.g., Statistics Canada methodologists). Funds should be set aside to cover these potential costs for more advanced research projects to ensure correct sample sizes calculations have been undertaken.

The successful implementation of the new process will require training for Epidemiology (e.g., sampling theory and methods) and team members (e.g., the new process for consulting Epidemiology). Epidemiology will need practical training to improve their abilities to obtain the

technical information required for sample size calculations from a non-technical audience without overwhelming staff or causing confusion.

Aside from the availability of Epidemiology Team resources, the main barrier to implementing the results of this Rapid Review is the perceived lack of confidence among Peel Public Health team staff related to numeracy. Some staff have expressed a general lack of confidence in using research and data to support public health practice. However, staff have a keen interest in addressing their knowledge gaps related to numeracy. Efforts are currently underway within the organization to improve staff confidence with numeracy, but these efforts have just begun and staff have not yet received the benefits of the proposed intervention(s). Staff may not yet see the need and/or benefit to engage in learning related to numeracy. Until the organizational culture of numeracy has improved, staff will need additional support from Epidemiology to conduct sound research. Until staff are self-sufficient in their use and understanding of numbers and basic statistics, the Epidemiology Team resources required to support teams will be substantial.

7 Decisions

As a result of this Rapid Review, the Epidemiology Team will:

1. Develop a process to consult with Peel Public Health staff for research projects, including sample size calculations. This iterative process will occur early on in the development of new projects and will carefully consider costs, time and resources available. This consultation process will include simple, standard questions to be answered by staff regarding their project (e.g., objectives, target population, outcomes of interest)
2. Calculate sample size estimates for probability samples using population-based or power analysis-based approach, depending on the goal of the project. Epidemiology will

identify the training required to conduct sample size calculations based on the power analysis approach.

3. Further examine the applicability of sample size calculations and statistics for non-probability samples. This will include providing staff with guidance on the use of results from convenience samples and other non-probability methods.
4. Identify external sources of support for sample size calculations (e.g., Public Health Ontario, researchers and consultants) and identify situations where these support services may be required.
5. Build an internal process for identifying knowledge gaps, training needs and structures for on-going learning (e.g., the existing **Learning and Brainstorming Sessions (LABS)** process).
6. Work with teams to increase their knowledge and capacity related to research methods and confidence with numeracy.
7. Work with middle management to engage teams in the new process for research consultations with Epidemiology. This process will require pilot testing of the new process and reflection back on lessons learned.
8. Encourage teams to implement where possible, a pilot phase for their research projects to refine sample size calculations and improve overall study quality.
9. Work with external sources of support to answer outstanding technical questions which were not adequately addressed by this Rapid Review.

Glossary of Key Terms

Alpha: (also called Type I error) is the probability of concluding that there is a significant difference when in fact there is no difference. The maximum probability of this error is "alpha", often chosen to be 5%.

***A priori* power analysis**: refers to power analysis conducted prior to the research study.

Beta: (also called Type II error) is the probability of concluding that there is no significant difference when in fact there really is a difference. The probability of this "beta" is the same as 1 minus the power of a study.

Cluster Sampling: enables random sampling from either a large population or one that is geographically diverse. Cluster sampling is often done to reduce costs by increasing sampling efficiency.[2]

Design effect: an adjustment used to account for a specific form of bias attributable to intraclass correlation in surveys using complex sampling design.[7] To calculate the design effect one must take the ratio of the sampling variance in the sampling method actually used to the sampling variance if SRS were used.[2]

Effect Size: the smallest difference or effect that the researcher considers to be clinically relevant. Determining the effect size can be challenging and is usually based on data from previous studies.

Elements: a unit (person, object) of a population.

Non-Probability Sampling: elements of the population have an unknown chance of being selected. Types of non-probability sampling include: convenience sampling, purposive sampling, quota sampling and snowball sampling.

No sampling error: also known as "bias". It is the statistical error caused by human error in the study design. This type of error may relate to: 1) selection bias, 2) non-response bias, 3) response bias. Non-sampling errors are extremely difficult, if not impossible, to measure. Bias caused by systematic errors cannot be reduced by increasing the sample size.

Population: the collection of elements from which a sample may be drawn.

Population parameter: The estimate of an outcome of interest if measuring the entire population. This value is often unknown and estimated using sample data.

Power (1-ß): the ability of a study to demonstrate an association if one exists. The power of a study is determined by several factors including the frequency of the condition under study, the magnitude of the effect, the study design, and sample size.

Probability Sampling: process of selecting elements in a sample such that each element has an equal chance of being selected. Probability sampling strategies allow estimates of sampling error to be calculated. Types of probability sampling include: simple random sampling, systematic random sampling, stratified sampling and cluster sampling.

Sample: the group of people whom you select to be in your study. It is a subset of the population elements that results from a sampling strategy.[2] The sample is intended to give results that are representative of the whole population.[7]

Sampling: the process of selecting units (e.g. people, schools, organizations) from a population of interest. Ideally, studying the sample will allow us to generalize our results back to the population from which they were chosen.[8]

Sampling error (SE): the part of the estimation error of a parameter caused by the random nature of the sample.[7] This type of error occurs by chance (too many elements of one kind and not enough of another). In general, an increased sample size is associated with decreased sampling error. The standard error will be small if the population is relatively homogeneous.

Sampling Frame: the listing of the accessible population (list, index, records) from which the sample will be drawn. This listing may not be totally inclusive of the study population.

Simple Random Sampling: "a commonly used simple random sampling procedure is to assign a number to each element in the sampling frame and use an unbiased process to select elements from the sampling frame." [2]

Stratified Sampling: the sampling frame is divided into non-overlapping groups or strata (e.g. sex, age groups). Then a random sample is taken from each stratum. This technique can be used to study a small subgroup of population that could be excluded using simple random sampling procedures.

Type I error: see "alpha" definition

Type II error: see "beta" definition

Table 1: Comparison of the Population-Based Approach and the Power Analysis Approach to Sample Size Determination

	Population-Based Approach	Power Analysis Approach
Definition	Approach use if interested in obtaining survey estimates that can be generalized back to a larger population.	Approach used when interested in detecting statistical differences between two groups (or sub-groups) of interest.
When Is it Used?	Used when interested in making inferences about a population of interest.	Used for research questions aimed at hypothesis testing and/or hypothesis generating.
Examples	1. A survey aimed to assess a variety of health behaviours among school-aged children. 2. A telephone survey aimed to explore food behaviours among Peel residents.	1. A study aimed to determine if a school-based physical activity intervention is effective in reducing BMI measures when compared to a group of students who have not received the intervention (control group). 2. A study aimed to explore sex differences in attitudes and behaviours related to diabetes prevention among South Asian populations in Peel.
Notes	The approach traditionally used to answer public health practice questions.	The approach most commonly seen in research. Should <u>not</u> be confused with the terms "a posteriori power calculation" "post hoc analysis", or "retrospective power analysis". In other words, it is not an approach used to determine if the final sample is adequately powered post analysis.

References

1. Statistics Canada. Sample size determination and allocation. In: Survey Methods and Practices. Canada: Minister of Industry; 2003. p. 151.

2. Dattalo P. Determining sample size: Balancing power, precision, and practicality. Oxford University Press, USA; 2008.

3. Levy PS, Lemeshow S. Sampling of populations: Methods and applications. 1999.

4. Lemeshow S, Hosmer D, Klar J, Lwanga S. Adequacy of sample size in health studies. Wiley New York; 1990.

5. Cohen J. Statistical power analysis for the behavioural sciences. Lawrence Erlbaum; 1988.

6. Kraemer HC, Thiemann S. How many subjects? Statistical power analysis in research. Sage Publications, Inc; 1987.

7. Last JM. A dictionary of epidemiology. Oxford University Press, USA; 2000.

8. Trochim WMK. Research methods knowledge base. Toronto: Atomic Dog Publishing; 2006.

Appendices

Appendix A1: Conceptual Map

Appendix A2: Analytical Map

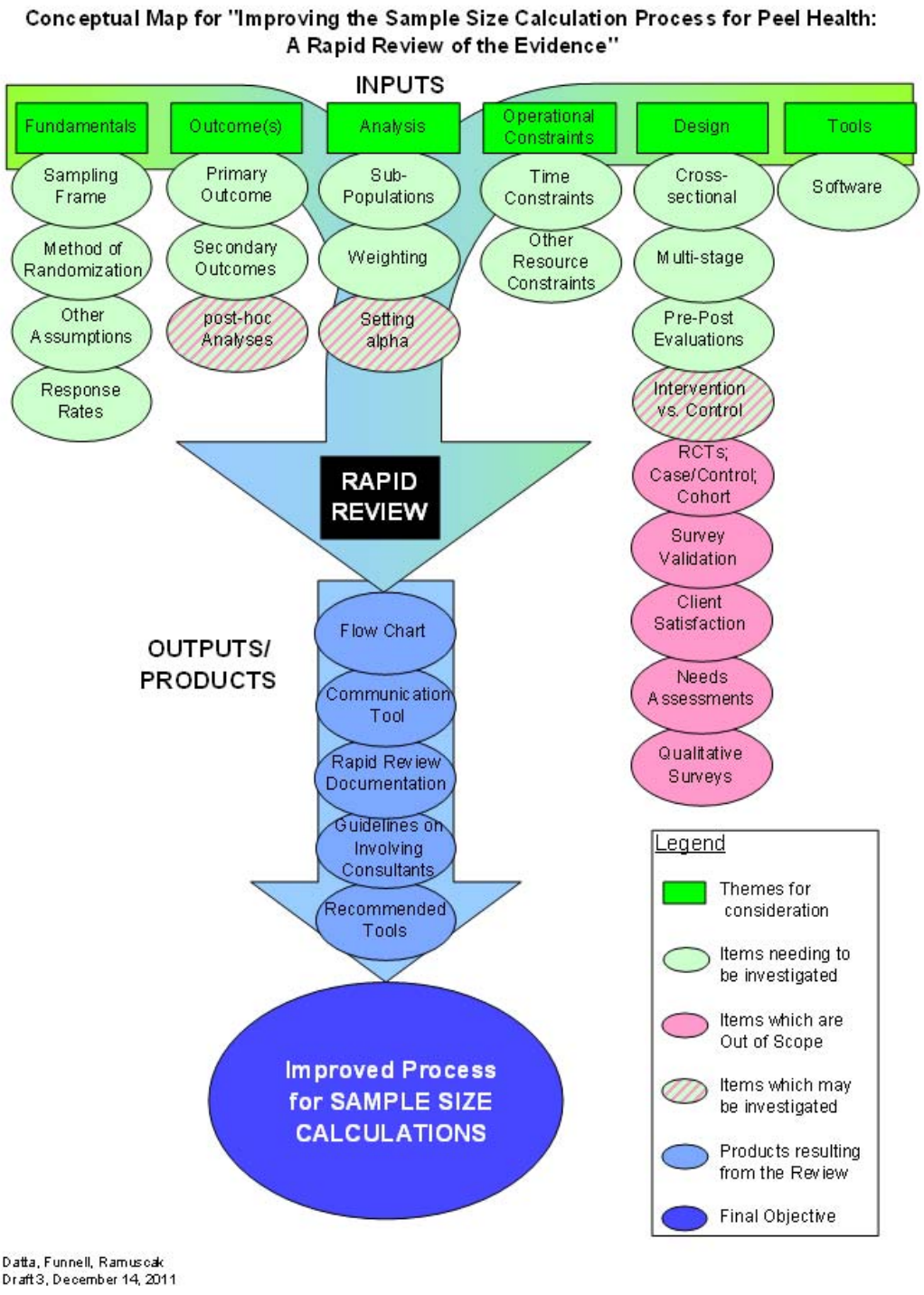
Appendix B: Overview of Search Process

Appendix C: Critical Appraisal Worksheet

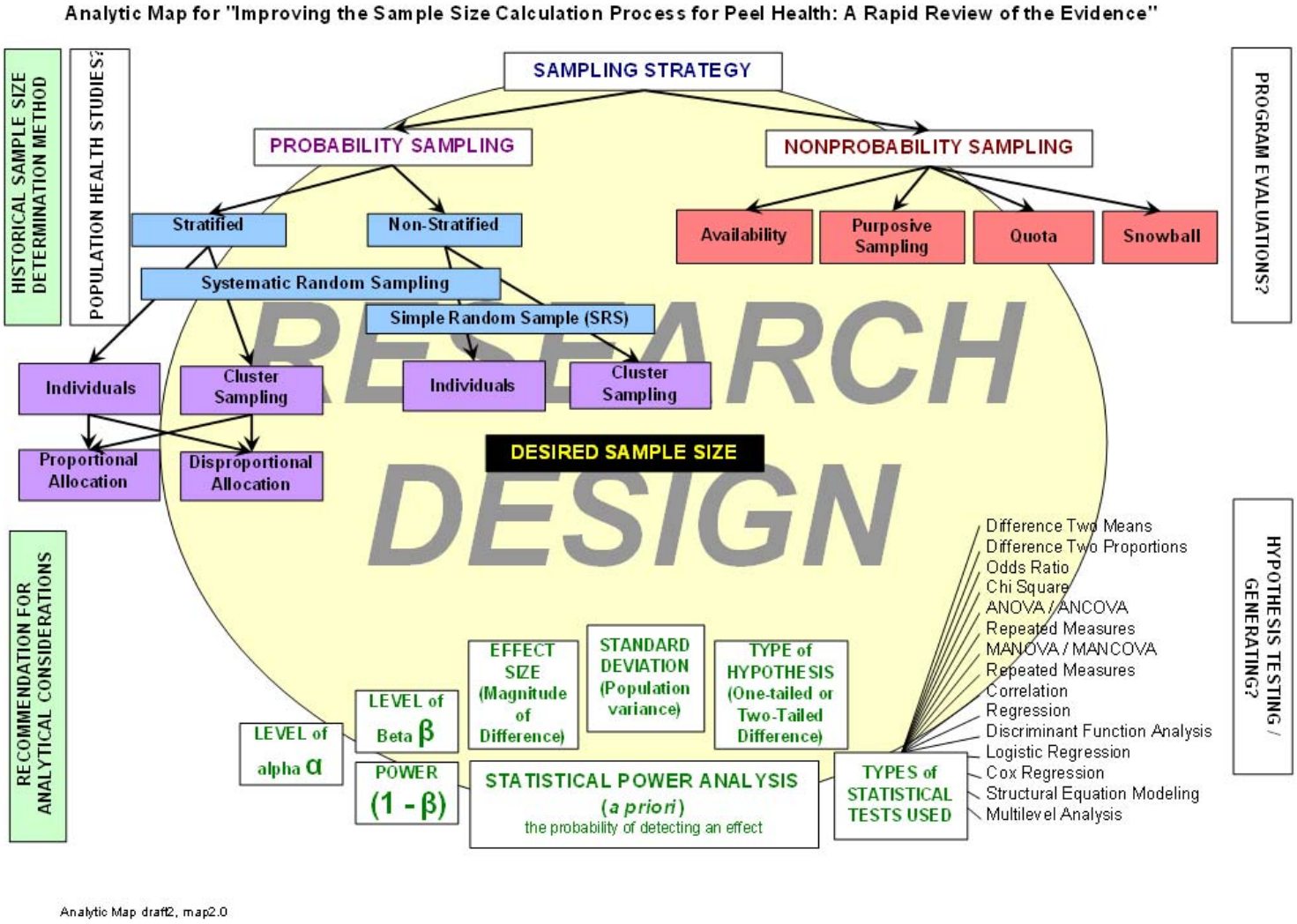
Appendix D: Data Extraction Tool for Textbook Comparison

Appendix E: Applicability & Transferability Worksheet

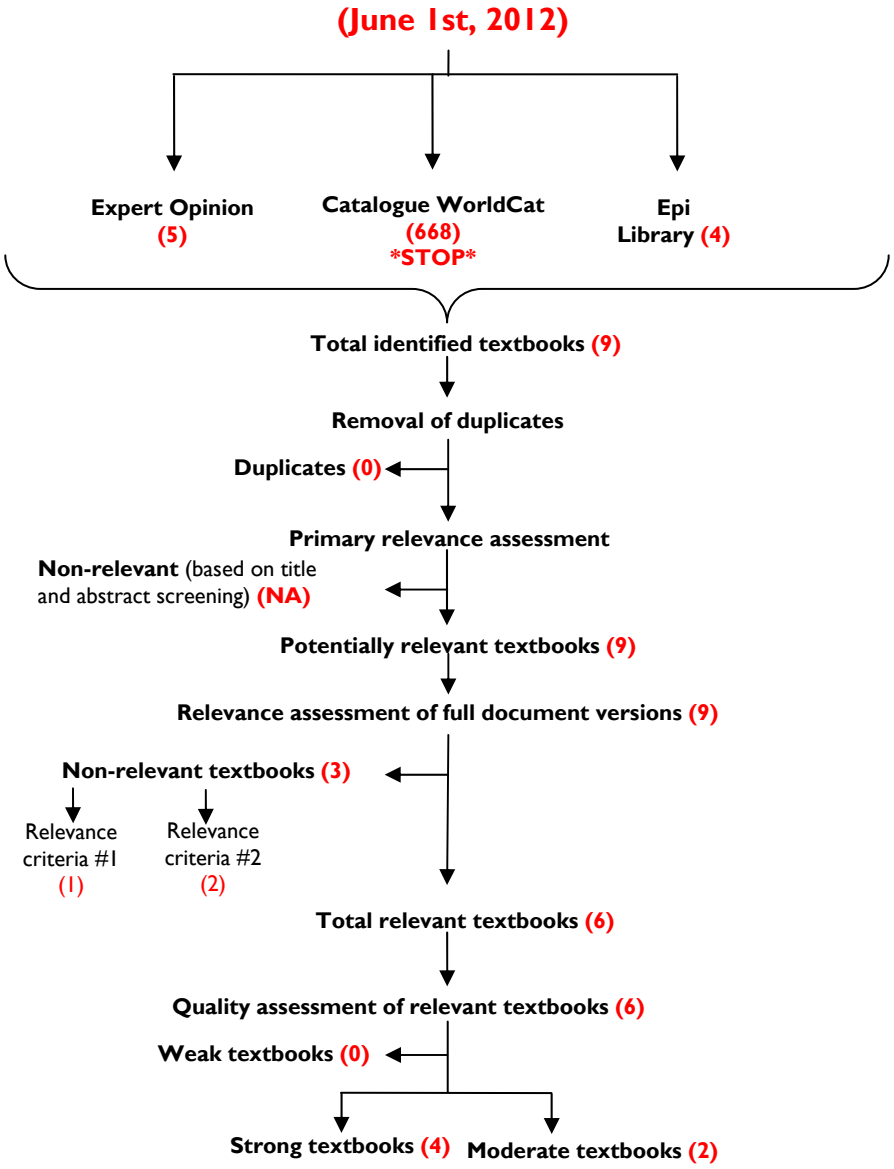
Appendix A1: Conceptual Map



Appendix A2: Analytical Map



Appendix B: Overview of Search Process



Source: Health-evidence.ca. (2009, November 25). *Keeping Track of Search Results: A Flowchart.* Retrieved [February 29, 2012]
Link to tool: http://www.health-evidence.ca/public/tools/10/Keeping_Track_of_Search_Results_-_A_Flowchart.ppt.

Notes:
Relevance Criteria #1: More recent edition of textbook available
Relevance Criteria #2: Does not provide sufficient detail on practice question

Summary of Search

Search Terms: phrase “sample size” searched in “Subject” field

Restrictions:

1. Books
2. English language
3. No year restrictions imposed

Search Results:

- 27 records at Ontario Universities through Scholars Portal RACER (Rapid Access to Collections by Electronic Requesting)
- 14 in Amicus, the Canadian National Catalogue
- 128 in LocatorPlus which is the catalogue of the collections of the National Library of Medicine
- 668 in WorldCat which is the world’s largest network of library content and services – most libraries that use OCLC (a world wide library services cooperative) for cataloguing automatically add their records to WorldCat

Appendix C: Critical Appraisal Worksheet

Critical Appraisal of Textbooks, Textbook Chapters
APPRAISAL WORKSHEET

Source: Survey Methods and Practices (Catalogue no. 12-587-X), Chapter 8 - Sample Size Determination and Allocation
Date of Appraisal: February 28, 2012
Completed by: Sheila Datta (R1); Nancy Ramuscak (R2); Karen Funnell (R3)

Criteria	Questions	Reviewer 1	Reviewer 2	Reviewer 3	Notes
1. Initial Appraisal Appraise a source by first examining the bibliographic citation – this can help determine the usefulness of the source. Bibliographic citations characteristically have three main components: author, title, and publication information					
Author - Consider the author's credentials, institutional affiliation, past writings and experience	Is the book/chapter written on a topic that is the author's area of expertise?	Y	Y	Y	Statistics Canada, Mandate to provide statistical information and analysis and promote sound statistical standards and practices http://www.statcan.gc.ca/about-apercu/mandate-mandat-eng.htm
	Have you seen the author's name cited in other sources or bibliographies?	Y	Y	Y	
	Is there a declaration of any conflict of interest for each author?	N	N	N	
Date of Publication	When was the source published?	2003	2003	2003	Original manuscript published October, 2003
	Is the source current (C) or out-of-date (O) for your topic?	C	C	C	
Edition or Revision - (Note: Further editions indicate a source has been revised and updated to reflect changes in knowledge, include omissions, and harmonize with its intended reader's needs). Many printings or editions may indicate that the work has become a standard source in the area and is reliable.	Is this a first edition of this publication?	N	N	N	Released September 2010 in electronic format. Adapted from materials used in StatsCan training, especially the *Survey Skills Development Manual* for the National Bureau of Statistics of China under auspices of the Canada-China Statistical Co-operation Program.
	Does the author discuss who reviewed previous editions or the process used to revise more recent editions? (e.g. in Preface, Foreword, or elsewhere in the text)	Y	Y	Y	
Publisher - If the source is published by a university press, it is likely to be scholarly. Although a source from a reputable publisher does not necessarily indicate quality, it does suggest that the publisher has a high regard for the source being published.	Is this a reputable publisher?	Y	Y	Y	Published by authority of the Minister responsible for Statistics Canada

Appendix C (continued)

Criteria	Questions	Reviewer 1	Reviewer 2	Reviewer 3	Notes
2. Content Analysis Examine the body of the source. Scan the table of contents and the index to obtain an overview of the material it covers. Read the chapters that specifically address your topic.					
Intended Audience - What type of audience is the author addressing?	Is the publication aimed at a specialized (S) or a general (G) audience?	S	S	S/G	Audience = Researchers, StatsCan Methodologists. "...main focus of manual is basic survey concepts useful to all readers, but some chapters are more technical... general reader may selectively study sections of technical chapters, choosing to skip more advanced material...Chapter 8, formulas used to determine sample size, requiring more technical understanding, begin with section 8.1.3 Sample Size Formulas."
	Is this source too elementary (E), too technical (T), too advanced (A), or just right (JR) for your needs?	JR	JR	JR	
Use of Bibliography - Note whether bibliographies are included. (The presence and quality of a bibliography at the end of a chapter may reflect the care with which the authors have prepared their work.)	Are bibliographies included?	Y	Y	Y	Most references are older (1962 to 1999). No Cohen citations. Does include Levy and Lemeshow citation. Also Hidiroglou, Kish, Särndal
	Are the references cited of high quality?	Y	Y	Y	
	Are the references current?	N	N	N	
	Do the references include work from the experts in the field or topic of interest?		N	Y	
Objective Reasoning	Is the information covered opinion (O) or fact (F)? (Facts can usually be verified, opinions evolve from the interpretation of facts).	F	F	F	Some opinion on methods but mostly Fact. Information is valid and well-researched.
	Is the author's point of view objective and impartial?	Y	Y	Y	
	Is the language free of emotion-arousing words and bias?	Y	Y	Y	
	Does the information appear to be valid (V) and well-researched or is it questionable (Q) and unsupported by evidence?	V	V	V	
	Are the ideas and arguments advanced more or less in line with other works you have read on the same topic?	Y	Y	Y	
Coverage - (Primary sources are the raw material of the research process. Secondary sources are based on primary sources)	Is the material primary (P) or secondary (S) in nature?	S	S	N/A	Is a compilation of other sources' information OR is Not Applicable as not research-based.
	If primary, does the work update other sources, substantiate other materials you have read, or add new information?	N/A	N/A	N/A	
Usefulness/Practicality	Does the content provide a complete description of the topic?	Y/N	Y/N	Y/N	Still unclear on allocation. Missing some details. Vague description of allocation. Power allocations referred elsewhere; only one cost allocation method is shared; multivariate allocation schemes referred elsewhere. Provides step by step approach pp.158-159 Provides examples (not necessarily PH)
	Does the author identify outstanding issues which are not covered?	Y/N	Y/N	Y/N	
	Does the textbook consider factors specific to the practical aspects of public health practice? (In other words, does it balance theory with practice?)	Y	Y	Y	
Writing Style	Is the publication organized logically?	Y	Y	Y	Very well organized from general content to more specific. Follows their stated content plan. Not repetitive except where necessary. Step-by-step.
	Are the main points clearly presented?	Y	Y	Y	
	Is the author's argument repetitive?	N	N	N	

Appendix C (continued)

Criteria	Questions	Reviewer 1	Reviewer 2	Reviewer 3	Notes
3. Evaluative Reviews Locate critical reviews of books in a reviewing source (such as Book Review Index, Book Review Digest or ProQuest Research Library)					
Published Evaluative Reviews	Is the review positive?				Could not find any reviews in a Google Scholar search. Book Review Index requires subscription. Book Review Digest requires subscription - applied for free trial through EBSCO ProQuest does not provide free trials to individuals.
	Does the reviewer identify other books that are more appropriate? If so, locate these sources for more information on your topic.				
	Do various reviewers agree on the value or attributes of the book or has it aroused controversy among the critics?				
Recommendation from Scholars and Peers	Is the book considered a valuable contribution to the field by practitioners and academics?			Y	University of Missouri: "There is a broad range of articles published in top journals such as Annals of Statistics, JASA, JRSSb, Biometrika, but the leading journal in the field dedicated solely to survey statistics is Survey Methodology published by Statistics Canada."
4. Citations Conduct a citation search on					
Citation Search (see next page)	How many other texts or articles cite the text or chapter?			6	Google Search found 6 citations and 3 university courses that use this text as either required or optional.
5. OVERALL RATING Based on the foregoing appraisal, provide an overall rating of this Textbook or Chapter (S = Strong; M = Moderate; W = Weak) and any rationale.					
Rating	How would you rate the text or chapter?	S	S	S	STRONG. 1. Source is Statistics Canada; their mandate is to give sound advice. 2. Audience is intended to be Methodologists training at statistical institutions. 3. Covers practical aspects in a 4-step approach, balancing practicality and precision. 4. Is written to be a How-To manual. 5. The manual has been used in other sectors and revised over time. 6. Is written in clear language and uses good examples.

Adapted with permission from: Olin Library Reference, Research & Learning Services, Cornell University Library, Ithaca, NY, USA

Appendix C (continued)

List of Citations or Relevant Uses

Citations

Statistics Canada Quality Guidelines, Catalogue no. 12-539-X, Fifth Edition, October 2009, Minister of Industry, 2009
Citations in References section of Chapters 1,3,4,6,11 and 14
Found on UN's Statistics web site: http://unstats.un.org/unsd/dnss/docs-nqaf/Canada-12-539-x2009001-eng.pdf

Lys P. The Current State of Cost of Production Estimates. A review of Country Practices in Ethiopia, Mali and Zambia. December 2010.
http://www.fao.org/fileadmin/templates/ess/ess_test_folder/Workshops_Events/production_cost/Lys_Report_on_country_case_studies_on_CoP.pdf

Boyko E and Hill C. An Introduction to Agriculture Statistics, April 2009, General Data Dissemination System, II World Bank
http://siteresources.worldbank.org/SCBEXTERNAL/Resources/Introduction_to_Agriculture_Statistics.pdf

Statistics Canada. Science Statistics - Industrial Research and Development, 2006 to 2010, Catalogue no. 88-001-X, December 2010 Edition
http://publications.gc.ca/collections/collection_2010/statcan/88-001-X/88-001-x2010006-eng.pdf

Körner T (Chair). Output Quality II - European Conference on Quality in Official Statistics, 2008
Hella H, Härkönen T, Salmela M. Statistical Quality Control: Pilot Experiments of Finnish Financial Survey Data
http://q2008.istat.it/sessions/30.pdf

College of Family Physicians of Canada (CFPC), Canadian Medical Association (CMA) and the Royal College of Physicians and Surgeons of Canada (RCPSC). 2007 National Physician Survey (NPS) Methodology

http://www.nationalphysiciansurvey.ca/nps/2007_Survey/pdf/2007.NPS.Methodology.and.Generalizability.of.Results_FINAL.pdf
http://www.ehealth-benchmarking.eu/results/documents/cases/eHBench_GP_CAN.pdf

University Course Material

University of Toronto Course "Survey Methods in Health Sciences II" CHL5204H
Required Textbook.
http://www.phs.utoronto.ca/courses/files/148_SURVEY_syllabus_2011_Jan.doc

University of Missouri Course "Stat 9100.3: Analysis of Complex Survey Data"
Recommended Reading
http://web.missouri.edu/~kolenikovs/Stat9100svy/Stat9100-003-ComplexSvy-syllabus.pdf

University of Ottawa Course "Research Methods for Public Policy"
Optional Textbook
http://ssms.socialsciences.uottawa.ca/vfs/.horde/offre_cours/syllabus/00101612126_API5136D.pdf

Appendix D: Data Extraction Tool for Textbook Comparison

Items Reviewed	Textbook/Chapter 1	Textbook/Chapter 2	Textbook/Chapter 3
General Information			
Author, Title, Year	Statistics Canada, Survey Methods and Practices 2003	Paul S. Levy and Stanley Lemeshow Sampling of Populations Methods and Applications 1999	Lemeshow, S., Hosmer, D.W., Klar, J. & Lwanga, S.K. Adequacy of Sample Size in Health Studies 1990
Overall Rating (from Appraisal step)	Strong	Strong	Strong
Intended Audience / Target Group	Researchers Stats Can Methodologists	Health professionals and students in Public Health graduate schools	Public Health professionals and Researchers
Details			
Number of chapters dedicated to topic	Chapter 8; p.358 (Case Study)	Portions of nine chapters, each based on particular type of survey design	Nine chapters based on particular type of analysis to be used and type of survey design
Objectives of Text	A practical guide to survey planning, design and implementation;	To describe the design and analysis of sample surveys	A practical guide to survey planning, design and implementation; Aims to address common challenges encountered in practice.
Which orientation? Population-based Approach or Power Analysis	Population-based Approach	Population-based Approach	Both Population-based Approach and Power Analysis
Sample Size Calculation Technique ("Intervention")			
What topics were well-covered by this text?	Uses a 4 steps approach balancing practicality with precision; Includes practical examples. Covers very well objective #1 (population estimates with multiple outcomes and sample size for clusters	Great amount of detail regarding sampling designs for objective #1 and #2	The application section includes a number of tables and formulae for the estimation for different types of analysis and different study designs.
What topics were not well-covered by this text?	Program Planning and Evaluation; Discussion on allocation (not clear)	Non-probability samples not addressed at all	Methods of sample size determination which do not use the normal approximation to the exact distribution are beyond the scope of this book. Does not discuss clusters, stratified sample design at all.
How well does this text step you through the process of sample size determination?	Provides clear steps for objective #1 and #2; Does not help with sample size determination for objective #3.	brief description of sample size calculation formulas for each type of sample design; illustrative examples of sample size determination; includes examples of the use of STATA and SUDAAN software	Provides brief description of sample size calculation formulas for each type of analysis or sample design; provides illustrative examples of sample size determination;

Appendix D (continued)

Items Reviewed	Textbook/Chapter 1	Textbook/Chapter 2	Textbook/Chapter 3
Practicality ("Results")			
What does this text suggest you do under scenario 1? (Population health survey where multiple outcomes are examined)	Recommends a 4-step approach: 1. Calculation of initial sample size, 2) Adjust for size of population, 3) Adjust for Design Effect, 4) Adjust for anticipated response rate	Gives formula for sample surveys of different types: simple random sampling and stratified sampling. Has illustrated examples of each situation.	Provides formulas for different study design. Uses power analysis to determine sample size. Must know a priori the types of analysis (and appropriate statistical test to be used).
What does this text suggest you do under scenario 2? (Population health survey where multiple outcomes are examined and where the sample selection involves clusters)	The design effect alters the number of cases needed. Proposes calculation of sample sizes under different scenarios: e.g.. 1. Setting outcome of interest at 50%; 2. calculation for main outcome of interest; 3. Determining what can reasonably be done and determining the power of the study.	Givens formula for calculating sample size for clusters using both one stage cluster sample and two-stage cluster samples (with and without equal probability of selection)	Cluster design discussed as a sampling strategy; Does not discuss how clusters and stratification affect sample size calculations.
What does this text suggest you do under scenario 3? (Consultation requests that involve program planning and evaluation)	Not discussed	Does not include sample size for non-probability samples	Does not include sample size for non-probability samples
Can OpenEpi be Used to Calculate the Scenario?			
Check against OpenEpi for scenario 1	Yes	Yes	Power calculation
Check against OpenEpi for scenario 2	Yes	? Yes with change to design effect	Power calculation
Check against OpenEpi for scenario 3	N/A	N/A	Power calculation
Concerns about when to use, when not to use ("Comments")			
Comments, strengths, concerns, limitations	Oversimplifies. Does not say what to do when certain assumptions are violated (e.g. When SRS is violated)	Gives a great deal of detail about the different types of methods that can be used. Good introductory text, although the sample size estimation section was fairly short. The text does outline a number of "preliminary planning" steps that should be conducted including: developing objectives, deciding on the population to be extrapolated to, identifying other potential methods for obtaining the information, identify the decisions to be made with the survey results, determine the required precision of the results, subgroups of interest, and available resources and time.	Uses a power analysis approach to sample size determination. Must have strong understanding of type of analysis to be done a priori (e.g. hypothesis testing and identifying appropriate test to use).

Appendix D (continued)

Items Reviewed	Textbook/Chapter 4	Textbook/Chapter 5	Textbook/Chapter 6
General Information			
Author, Title, Year	Jacob Cohen Statistical Power Analysis for the Behavioural Sciences (Second Edition) - 1988	Patrick Dattalo Determining Sample Size - Balancing Power, Precision and Practicality 2008	Kraemer, H.C., Thiemann, S.; How Many Subjects? Statistical Power Analysis in Research. 1987
Overall Rating (from Appraisal step)	Moderate	Moderate to Strong	Strong
Intended Audience / Target Group	States that the text has an "Applied viewpoint" but really this text is aimed at behavioural researchers	Researchers with knowledge of statistics	Public Health professionals and Researchers
Details			
Number of chapters dedicated to topic	Portions of each chapter (sample size tables for each type of test)	All six - entire volume	Chapter 2 introduces a single method appropriate for calculating power. The rest of the textbook aims to statistically prove the rigor of this method under different scenarios.
Objectives of Text	Comprehensive examination of power analysis for researchers - based on the type of analytical test being conducted	Presents approaches to determining sample size: Power analysis, Confidence Intervals and Computer-Intensive Strategies	Provides a practical guide to sample size calculations for a wide variety of research designs.
Which orientation? Population-based Approach or Power Analysis	Power Analysis	Power Analysis	Power Analysis
Sample Size Calculation Technique ("Intervention")			
What topics were well-covered by this text?	Power analysis for various types of statistical tests: t tests for means, comparison of proportions, correlations, chi-squares, regressions, ANOVA and ANCOVA	Basic Terms and Concepts (p3-8;11-12) Statistical Power (p13-18) Confidence Intervals (p 33-42) many examples of statistical tests many suggestions & links to software, tools	Introduces a single method appropriate for calculating sample size and study power with only slight modifications for different statistical tests.
What topics were not well-covered by this text?	The type of study design required for the various types of tests. Doesn't discuss clusters, stratified sampling, study design at all.	No discussion of study design approach Nonrandom / Nonprobability discussed briefly without useful solution	The type of study design required for the various types of tests. Doesn't discuss clusters, stratified sampling study design at all.
How well does this text step you through the process of sample size determination?	Has formula needed to calculate parameters to use power tables and sample size tables included in each chapter. Steps are not clearly defined.	Different Methodology suggested: 1. Establish a desired sample size based on feasibility &/or other studies 2. Determine alpha e.g. .05 3. Calculate beta assuming alt Hypothesis is true 4. Calculate power (1 - beta) 5. Examine resulting power and compare to .80	Single method simplifies the process for calculating sample sizes. Emphasis on calculating power prior to the initiation of a study.

Appendix D (continued)

Items Reviewed	Textbook/Chapter 4	Textbook/Chapter 5	Textbook/Chapter 6
Practicality ("Results")			
What does this text suggest you do under scenario 1? (Population health survey where multiple outcomes are examined)	Method depends on the type of analytical test to be completed in this scenario.	Not really addressed - uses hypothesis testing/generating approach Focus on most important research question Determine analysis strategies Apply steps 1-5 as above	Uses one table to estimate sample size or power in the various test situations. Tries to eliminate most of the mathematical complexity of power calculations. Requires the user to know 2 of the 3 parameters needed to determine the sample size using the tables provided.
What does this text suggest you do under scenario 2? (Population health survey where multiple outcomes are examined and where the sample selection involves clusters)	??? Does not address clusters or design issues at all.	Minimal mention of clusters related to multi-level analyses (p35-36) "When cluster sampling is part of the design, a reasonable approach is to use a design effect of 2 or 3". "...a large number of groups is more important than a large number of individuals per group" Determine power for an SRS Adjust for design effect.	Does not discuss how clusters and stratification affect sample size calculations.
What does this text suggest you do under scenario 3? (Consultation requests that involve program planning and evaluation)	Method depends on the type of analytical test to be completed in this scenario.	Nothing specific, but might be similar to scenario 1. For nonrandom samples, (p 74): "Kruglanski (1975) argued that multiple studies aimed at investigating interrelated aspects of a phenomenon using different samples might generate more confidence in any single study's findings...combing the findings from a reasonably large set of nonprobability samples might be an acceptable, practical compromise."	Does not include sample size for non-probability samples
Can OpenEpi be Used to Calculate the Scenario?			
Check against OpenEpi for scenario 1	Power calculation	Power calculation	Power calculation
Check against OpenEpi for scenario 2	Power calculation	Power calculation	Power calculation
Check against OpenEpi for scenario 3	Power calculation	Power calculation	Power calculation
Concerns about when to use, when not to use ("Comments")			
Comments, strengths, concerns, limitations	Good background text related to power analysis. Highlights to importance of power analysis in ensuring that studies are adequately powered to detect differences that may exist between groups. The text does not examine the study designs required to collect the data for the significance test - does not address clusters, strata etc. The text is foundational but does not talk about the software that is now available to conduct power analysis.	Good for Analysis approach - many examples for various statistical tests Good for links to resources, references and tools Different methodology described, especially Computer-Intensive Strategies that suggest building model(s) to determine sample size.	determine sample size. Can be easily applied to practice but practitioner must have strong understanding of type of analysis to be done a priori (e.g. hypothesis testing and identifying appropriate test to use); Parameter estimates must be informed by the literature.

Appendix E: Applicability & Transferability Worksheet

Starting a New Program
Applicability and Transferability Worksheet

Factors	Questions	Notes
Applicability (feasibility)		
Political acceptability or leverage	• Will the intervention be allowed or supported in current political climate? • What will the public relations impact be for local government? • Will this program enhance the stature of the organization? o *For example, are there reasons to do the program that relate to increasing the profile and/or creative a positive image of public health*? • Will the public and target groups accept and support the intervention in its current format?	• Region of Peel CAO supports EIDM and the use of data • Teams will support this as long as the study can be completed • Peel Public Health wants to spend their money wisely and to ensure that the data is representative of the population (i.e. has enough sample to answer the research question) • This process asks the teams to do more work in thinking through and describing the methods of their project and how the data will be used (i.e. more iterative process than currently) • Conducting sound research and evaluation will increase our profile internally (within our organization and across the corporation) and externally (e.g., with PHO, other health units) • This process will be implemented within a larger process for the Research Review Committee • More onus will be placed more on teams to control their projects and therefore increase their ownership of the results
Social acceptability	• Will the target population find the intervention socially acceptable? Is it ethical? o *Consider how the program would be perceived by the population.*	• Teams want to do their research projects well but they are under time pressures (i.e., surplus funds must be spent between September and December). Epi should work

	o *Consider the language and tone of the key messages.* o *Consider any assumptions you might have made about the population. Are they supported by the literature?* o *Consider the impact of your program and key messages on non-target groups.*	proactively with the teams to identify these pressure points and plan early on how to spend funds during these times. Epi should stage the conversations about priority research questions early in the year (January) so that there is adequate time to identify the project to be funded, determine the best methods and problem solve issues when required. • Work with middle managers (Supervisors and Managers) to engage the teams in the new process. This will include working with Bev Bryant and the workforce development plans. • Epi should pilot test the new process with a small number of teams and reflect back on what went well and what needs to be improved. Be opportunistic in testing the new process with teams that are ready. Don't oversell our services until we are able to provide them. • Advise PHMT and get the Directors on board with the new process. Then work with middle managers to engage the teams. • Tie the new process to the larger changes required for the Research Review Committee and the ethics of conducting poor research. We need to emphasize to the teams the importance of good methods.

Available essential resources (personnel and financial)	• Who/what is available/essential for the local implementation?
	• Are they adequately trained? If not, is training available and affordable?
	• What is needed to tailor the intervention locally?
	• What are the full costs?
	  o *Consider: in-kind staffing, supplies, systems, space requirements for staff, training, and technology/administrative supports.*
	• Are the incremental health benefits worth the costs of the intervention?
	  o *Consider any available cost-benefit analyses that could help gauge the health benefits of the intervention.*
	  o *Consider the cost of the program relative to the number of people that benefit/receive the intervention.*
		• This new process feeds into the scope of the other work that Epi does with the teams related to research methods and design.
		• This is one piece of a larger picture. As this process rolls out, there will be an increase in demand for Epi services.
		• Teams will require assistance with implementation of the required sample size and problem solving issues that arise during data collection (e.g., inability to recruit a large enough sample)
		• This new process will require many more resources for survey design
		• Epi must be careful about how many resources they have to offer
		• Epi staff have other work priorities that must be balanced with research work – health status reports, Epi projects, analysis for the website
		• Support for ad-hoc requests from teams are not identified in the annual workplan processes or the time required for the support ends up being more than anticipated
		• Epi should make use of researchers and resources available from others (e.g., PHO, Stats Canada)
		• Peel Public Health should set aside funds for external consultations for more difficult projects to work through the methods
		• Epi needs to improve their own skills for facilitating and mentoring others to help teams (e.g., how to ask for technical information in a non-technical way and not overwhelm teams)
		• Identify sources of support for consultations when needed – either free services (PHO or

		researchers funded to provide support to community organizations) or for a fee (consultants) • Don't oversell the services that can be provided • Start small and roll out to select teams. Build in time to reflect back on what worked well and how things can be improved. • Must have dedicated time to support teams and learn from the process being piloted with teams • Must make time for our own team learning and training. • In general the cost of the resources put into the new process will be worth it • Epi team will need to identify the priorities for services provided and the situations where sample size calculations should be conducted by others
Organizational expertise and capacity	• Is the intervention to be offered in line with Peel Public Health's 10-Year Strategic Plan (i.e., 2009-2019, 'Staying Ahead of the Curve')? • Does the intervention conform to existing legislation or regulations (either local or provincial)? • Does the intervention overlap with existing programs or is it symbiotic (i.e., both internally and externally)? • Does the intervention lend itself to cross-departmental/divisional collaboration? • Any organizational barriers/structural issues or approval processes to be addressed? • Is the organization motivated (learning organization)? o *Consider organizational capacity/readiness and internal supports for staff learning.*	• Support for sample size calculations and research in general falls within the Surveillance strategic priority • Population health assessment is part of the Foundational Standard of the Ontario Public Health Standards • Over time, the new process will lead to a building of our knowledge and the knowledge of the teams. This will ultimately lead to better programs, decisions and the use of resources. • Need to find opportunities for training when they present themselves • Build into the process a second reviewer of sample size calculations when completed in-house • Middle management will be the focus of the change management related to this new

		process • Need to build in time for Epi team learning on the new process • Must acknowledge mistakes made in the past and learn from them. Epi team has to recognize our current level of understanding and identify our knowledge gaps. • Learning conducted to fill knowledge gaps will lead to an increased understanding of sample size calculations and research. This new knowledge can be applied to all new work projects.

Transferability (generalizability)		
Magnitude of health issue in local setting	• What is the baseline prevalence of the health issue locally? • What is the difference in prevalence of the health issue (risk status) between study and local settings? o *Consider the Comprehensive Health Status Report, and related epidemiological reports.*	• Currently conducting five to ten sample size calculations per year • Anticipate that this will increase with Research Review Committee process and increased expectations on teams for EIDM and research.
Magnitude of the "reach" and cost effectiveness of the intervention above	• Will the intervention appropriately reach the priority population(s)? o What will be the coverage of the priority population(s)?	• Epi must be careful about how many resources they have to offer – there are many teams that may require services – "don't get swept up in the enthusiasm".
Target population characteristics	• Are they comparable to the study population? • Will any difference in characteristics (e.g., ethnicity, socio-demographic variables, number of persons affected) impact intervention effectiveness locally? o *Consider if there are any important differences between the studies and the population in Peel (i.e., consider demographic, behavioural and other contextual factors).*	• Support for numeracy falls on the Epi team at this point in our organizational history. Epi needs to help teams to make sense of "the numbers" until such time as they are more confident and self-sufficient in the use of data and statistics.

Proposed Direction (after considering the above factors):

- Develop a process to consult with Peel Public Health teams on research projects that require sample size calculations.
- Conduct sample size calculations for research questions involving probability samples using either population-based or power analysis approach (where appropriate).
- Identify training opportunities for the epidemiology team to build skills and competency related to sample size calculations with particular emphasis on sample size calculations using the power analysis approach.
- Provide staff with guidance on the use of results obtained from convenient samples and non-probability samples.
- Identify external sources of support for sample size calculations that fall outside the expertise of the epidemiology team.
- Develop process for identifying knowledge gaps, training and on-going learning for health department staff and the epidemiology team.
- Develop a change management strategy to engage teams in the new process for research consultations with Epidemiology.
- Work with external sources of support to address technical questions on sample size calculations that have not been addressed through this Rapid Review.