Chapter 4

Methodology

The present work focuses on developing a new methodology to integrate the findings from studies of complex public health interventions. It involved a qualitative study on stakeholders’ perspective about complexity of public health interventions, development of a tool to quantify the complexity in public health interventions into a numerical score, assessment of validity and reliability of the tool, evaluation of statistical properties of the complexity score, classification of studies on the basis of their complexity status and adjusting the measured complexity in meta-analysis by a suitable statistical technique. This chapter elucidates the detailed methodology specific to each of the aforesaid objectives of the study.

4.1 Methodology - First objective

To identify the methodological challenges, particularly the complexities in the Population, Intervention, Context and Outcome in integrating the findings of studies pertaining to public health interventions

A qualitative study entitled “Complexity in Public Health Interventions - Stakeholders’ Perspective: A Qualitative Analysis”\textsuperscript{16} was carried out to probe into various aspects of complexity and excerpt views of several people associated with
the public health domain in an effort to bring more clarity towards understanding the term “complexity in public health interventions”. A combination of semi-structured one-to-one interviews and focus group discussions were incorporated. Public health professionals, who are closely associated with the discipline of public health either as public health stakeholders, students, researchers, faculty, statisticians and systematic reviewers were identified mainly through academic contacts (i.e., the people in the professional network) and were requested to participate in the study. Sessions were conducted in English language at a convenient place and time as suggested by the interviewees. A verbal informed consent was obtained and the conversations were audio-recorded. No incentive was provided for participation. The participants were queried on their perception about complexity in public health interventions. As the study was of exploratory type, the interview guide consisted of only a set of three questions (Text box 4.1). However, additional questions were posed during the sessions to seek further details. It was decided to continue with the sessions till information saturation (no emergence of new information or repeated occurrence of same information) was attained. The interviews/focus group discussions involved a moderator and a co-facilitator. Role of the moderator was to involve in conversation, where as co-facilitator was assigned the responsibility of audio-recording the conversation and summarizing entire discussion.

**Textbox 4.1**

**Interview guide**

1. According to you what is complexity in public health interventions?
2. What are the causes for complexity in public health interventions?
3. Do you think that complexity in public health interventions is a major issue?
4.1.1 Textual data analysis

Views shared by participants were transcribed into a document. Broad divisions of the perceptions were made based on opinions or ideas. The five sources of complexity as depicted in MRC guidance\textsuperscript{47} namely (1) The number of interacting components within the experimental and control interventions (2) The number and difficulty of behaviours required by those delivering or receiving the intervention (3) Number of groups or organizational levels targeted by the intervention (4) Number and variability of outcomes and (5) Degree of flexibility permitted in the tailoring of the intervention were regarded as themes for the consolidation of perceptions. This approach of qualitative data synthesis where the themes are derived from literature is referred to as ‘Deductive approach’.\textsuperscript{71} Particular codes were assigned for themes to facilitate easy consolidation of responses and coding was performed manually.

4.1.2 Conceptual framework for complexity in public health interventions

A concept map was created to provide a visual overview of the factors that bring about complexity in public health interventions and further depict the relationship between different factors.

4.2 Methodology - Second objective

To develop a suitable methodology to integrate the findings from several studies of Public Health interventions after accounting for the complexities

This objective involved four main activities namely; (1) Development of a tool
to measure complexity in public health interventions (2) Exploration of statistical properties of complexity score (3) Classification of studies based on their complexity status (4) Adjusting the measured complexity in meta-analysis using a suitable statistical technique.

4.2.1 Tool to measure complexity in public health interventions

The main purpose of the tool was to assess complexity in four major domains namely population, intervention, context and outcome. The qualitative study “Complexity in Public Health Interventions - Stakeholders’ Perspective: A Qualitative Analysis” provided a clear understanding of complexity in public health interventions. In addition, it contributed to the development of theories which describe the factors that contribute complexity in each of the four domains. With these theories as the background, the tool development activity began with the preparation of an initial checklist of items. Studies/articles those were included in the published public health systematic reviews were considered as the primary source of items for the tool. Therefore, 26 studies coming from six diverse Cochrane public health systematic reviews were retrieved. In the next step, each of the collected articles were thoroughly scrutinized and concisely described by the characteristics of four identified domains. From these summaries, an initial list of items was prepared by culling out the items reflecting complexity separately for the four domains in accordance to the deduced theories. Later, with much deliberations, discussions and consultation of methodological experts, relevant items were filtered from the initial list to have a final set of items. The scoring pattern for each item was developed on the basis of two criteria namely relative importance of the item and number of options in the item. The scoring was set in direct proportion to complexity i.e., a higher score represents a higher complexity. Finally, the
scores of all items of a domain were added to obtain the domain specific complexity score and four domain specific complexity scores were added to obtain the total complexity score. The developed tool was piloted on the collected 26 articles. Reliability of the tool was evaluated by computing Intraclass Correlation Coefficient (ICC) for the complexity scores of a set of 10 studies assessed for complexity using the tool independently by two raters. ICC was calculated for the total score as well as for score of each domain and eventually the first version of the tool was developed. ICC is given by

$$ICC^2(A,1) = \frac{MSS_{BS} - MSS_E}{MSS_{BS} + (k - 1)MSS_E + k/n(MSS_R - MSS_E)}$$

where,

- ICC2 represents class 2 of ICC in which all studies are evaluated all raters
- A represents absolute agreement between the raters
- 1 represents that all studies are evaluated by every rater only once
- k is the number of raters
- n stands for number of studies
- MSS_{BS} stands for mean sums of square between studies
- MSS_{R} stands for mean sums of square between raters
- MSS_{E} stands for mean sums of square due to error

In the next stage, the tool was upgraded by adding new items, refining existing items based on the experience from piloting process. Further, the prototype tool was reviewed by seven experts in the domain of public health for the purpose of content validation. The experts were requested to rate each item of the tool as either highly relevant, relevant, irrelevant or highly irrelevant and to suggest suitable modifications (if required). Items for which at least four experts (approximately half) rated as highly relevant were retained with no changes, whereas those rated
as relevant by at least four of them were retained with some minor modifications. Further, the items being marked as irrelevant by at least four experts, were subjected to a major modification and finally those marked as highly irrelevant by at least four of the experts were discarded. Modifications and refinements suggested by the experts were incorporated and the tool was finalized with five items under population domain, seven items under each of the intervention and context domain and four items under outcome domain. Content validation ensured that the tool does not miss out any potentially relevant items.

4.2.1.1 Reliability of the final tool

Number of studies to be scored to assess reliability was determined by the sample size formula put forward by Walter et al in their article on “Sample size and optimal designs for reliability studies”. For the null hypothesis $\rho_1 = \rho_0$ and alternate hypothesis $\rho_1 > \rho_0$ where, $\rho_0$ is the minimally acceptable reliability and $\rho_1$ is the specific underlying value of reliability under alternate hypothesis, the number of studies required is:

$$n = 1 + \frac{[Z_{1-\alpha/2} + Z_{1-\beta}]^2 k}{(lnC_0)^2(k - 1)}$$

$$C_0 = \frac{1 + k\theta_0}{1 + k\theta}$$

$$\theta_0 = \frac{\rho_0}{1 - \rho_0} \quad \& \quad \theta = \frac{\rho_1}{1 - \rho_1}$$

Where, ‘n’ is the number of studies required to be scored, ‘k’ is the number of raters, $Z_{1-\alpha/2} = 1.96$ at $\alpha = 0.05$ and $Z_{1-\beta} = 0.84$ at 80% power.

For a $\rho_0$ of 0.7, $\rho_1$ of 0.9 and three raters, number of studies to be assessed for complexity was found to be 16. However a set of 20 studies were assessed independently by three raters.
4.2.1.2 **Electronic version of the tool to measure complexity in public health interventions**

Electronic version of the developed tool was prepared using “Microsoft excel 2007 VBA (Visual Basic for Application) and Macros”. A total of five user forms were created; four corresponding to each of the four domains and one user form to compute the total complexity score by adding the four domain specific complexity scores. Every user form consists of all the items belonging to the respective domain. Provision was made to enter the score corresponding to each of the item and eventually obtain the sum of all the entered scores.

4.2.2 **Statistical properties of the complexity score**

Distribution fitting - determining the ideal or best probability distribution that fits the total complexity score comprised following three steps;

1. Anticipating the possible distributions that could fit the total complexity score: the distributions that were likely to fit the complexity score were identified by constructing a histogram out of total complexity score of 259 studies from 30 published systematic reviews of public health importance. Based on the shape of histogram, a possible list of distributions was generated.

2. Estimating the parameters of the distributions: the parameters of the anticipated distributions were estimated by maximum likelihood technique.

3. Evaluating the quality of fit: Anderson-Darling goodness of fit test was used to evaluate the quality of fit i.e., to decide which among the anticipated distributions was the best fit. Higher p-value (>0.05)/lower test statistic value was the criteria to decide the best fit.
4.2.3 Classification of studies based on their complexity status

The total complexity score is obtained by addition of complexity scores of four domains namely population, intervention, context and outcome. Number of items, minimum attainable and maximum attainable scores are not same in each of the four domains. For instance, a study can have high total complexity score which might be contributed by any one of the four domains, while other domains might possess relatively low scores. Therefore, it is important to take into account the relative contribution of all the four domains towards the total complexity score in ranking the studies according to their complexity status. The technique of composite dynamic index developed by NS Iyengar and P Sudarshan\textsuperscript{364} was adopted to rank the studies as per their complexity status and further classify the studies into four classes of complexity based on their complexity score namely very highly complex, highly complex, moderately complex and least complex for the purpose of interpretation.

The detailed step-by-step procedure is elucidated below;

**Step I: Transforming the complexity scores**

The domain specific complexity scores were subjected to following transformation;

\[
X_{ij} = \frac{Y_{ij} - \text{Min} Y_{ij}}{\text{Max} Y_{ij} - \text{Min} Y_{ij}}
\]

(4.3)

\[
i = 1,2,3,4 \text{ domains, } j = 1,2,3...,n \text{ studies}
\]

\[Y_{ij}\] represents the observation i.e., complexity score in the \text{i}^{\text{th}} domain and \text{j}^{\text{th}} study. Domain 1 - population, domain 2 - intervention, domain 3 - context and domain 4 - outcome. For instance, \(Y_{13} = 4\) means that the complexity score of population
domain of the third study is 4.

Min \( Y_{ij} \) and Max \( Y_{ij} \) represent respectively the minimum and maximum of \( Y_{i1} , Y_{i2}, . . . \) \( Y_{in} \). \( X_{ij} \) represents the transformed complexity scores, which are also referred to as standardized scores or scaled scores and they range between 0 and 1.

The main purpose of this transformation or standardization procedure is to achieve equality in the score variability across all the four domains.

**Step II: Assigning weights to the standardized complexity scores**

As the relative contribution of four domains towards total complexity score is not equal, it is necessary to assign weights to each of the domains. The weights were assigned inversely proportional to the variation in the values of respective domains.

\[
W_i = \frac{H}{SD(X_i)}, \quad \text{where} \quad H = \left[ \sum_{i=1}^{4} \frac{1}{SD(X_i)} \right]^{-1} \quad (4.4)
\]

\[\therefore W_i \propto \frac{1}{SD(X_i)}\]

\( W_i \in (0, 1) \) and \( \sum_{i=1}^{4} W_i = 1 \)

**Step III: Construction of the index**

The index was constructed as a weighted linear combination of the standardized complexity scores of four domains, given by:

\[
\bar{X}_j = W_1X_{1j} + W_2X_{2j} + W_3X_{3j} + W_4X_{4j} \quad (4.5)
\]

\[0 < \bar{X}_j < 1\]

Later the index values were arranged in the descending order of their magnitude to ascertain the relative rank of each study in terms of complexity.
Step IV: to classify the studies into four classes of complexity namely very highly complex, highly complex, moderately complex and least complex

Beta distribution (which is positively skewed and the ranges between 0 and 1) was fitted to the values of the developed index and studies were classified into four classes of complexity based on a “fractile classification”.

Probability distribution of a random variable ‘T’ following Beta distribution with parameters ‘p’ and ‘q’ is given by;

\[ f(T) = \frac{T^{p-1}(1 - T)^{q-1}}{\beta(p, q)} \]  \hspace{1cm} (4.6)

where, \( \beta(p, q) = \int_0^1 T^{p-1}(1 - T)^{q-1}dT \)

\[ 0 < T < 1 \text{ and } 0 < p < \infty, \, 0 < q < \infty \]

The values of ‘p’ and ‘q’ were estimated by maximum likelihood estimation method in ‘R’ software using “fitdisr” package, specifying the intial value of parameters p and q equal to unity (shape1 = 1 & shape2 = 1). Estimated values of ‘p’ and ‘q’ were used to find the cutoff points \( T_1, T_2 \) and \( T_3 \) such that \((0, T_1), \, (T_1, T_2), \, (T_2, T_3) \) and \((T_3, 1)\) represent the intervals covering \( \bar{X}_j \) with 25% probability each. These intervals were used as a basis for classification of studies into four classes of complexity namely very highly complex, highly complex, moderately complex and least complex, which is as given below;

1. Least complex if \( 0 < \bar{X}_j \leq T_1 \)
2. Moderately complex if \( T_1 < \bar{X}_j \leq T_2 \)
3. Highly complex if \( T_2 < \bar{X}_j \leq T_3 \)
4. Very highly complex if \( T_3 < \bar{X}_j \leq 1 \)
This technique was applied to scores of 259 studies which were assessed for complexity using the developed tool to rank them according to their complexity status as well as to classify them into four classes of complexity.

4.2.4 Adjusting the measured complexity in meta-analysis

Total complexity score was adjusted in meta-analysis using a technique known as ‘Meta-regression’. Meta-regression is one of the popular methods to access impact of covariates on the effect estimates of studies to be meta-analyzed. It helps to assess the relationship between the dependent variable, which is the effect estimate of studies (log odds ratio, log risk ratio or mean difference) and one or more study level covariates.\(^ {365,366}\) The fundamental difference between the meta-regression and simple linear/multiple linear regression is that in the former the variables are measured at the study level, whereas in the latter the variables are measured at the level of individuals.\(^ {8}\) Further, the meta-regression can also be used to obtain the pooled estimate after adjusting for the effect of covariates.\(^ {367}\)

In the present study, the total complexity score was considered as a covariate and adjusted in meta-analysis and corresponding pooled estimate adjusted for complexity was obtained. Detailed computational procedure of meta-regression involving a single covariate is illustrated below.\(^ {368}\)

4.2.4.1 Notations and definitions of meta-regression

Just as in meta-analysis, there are two approaches to meta-regression - fixed effect model and random effects model.

Meta-regression: Fixed effect

\[
Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \quad i = 1, 2, 3, \ldots, n \text{ studies} \tag{4.7}
\]
Where, $Y_i$ is effect estimate of $i^{th}$ study, $\beta$s are the regression coefficients to be estimated and $X_i$ is the value of covariate for $i^{th}$ study. Coefficients of the above model are estimated by “Weighted Least Squares” technique, defining weights by the reciprocal of variance of corresponding effect estimate i.e., $W_i = 1/\sigma_i^2$; $W_i$ is the weight and $\sigma_i^2$ is the variance of effect estimate of $i^{th}$ study. However standard error of the estimated regression coefficients needs to be subjected to a minor correction.\cite{365,369}

**Weighted least squares algorithm**

In presence of a single covariate, the regression coefficients are obtained by minimizing;

$$S = \sum_{i=1}^{n} W_i(Y_i - \beta_0 - \beta_1X_i)^2$$

(4.8)

After solving, it gives;

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} W_i \sum_{i=1}^{n} W_iX_iY_i - \sum_{i=1}^{n} W_iX_i \sum_{i=1}^{n} W_iY_i}{\sum_{i=1}^{n} W_i \sum_{i=1}^{n} W_iX_i^2 - (\sum_{i=1}^{n} W_iX_i)^2}$$

$$\hat{\beta}_0 = \bar{Y}_W - \hat{\beta}_1\bar{X}_W$$

Where, $\bar{Y}_W = \frac{\sum_{i=1}^{n} W_iY_i}{\sum_{i=1}^{n} W_i}$ and $\bar{X}_W = \frac{\sum_{i=1}^{n} W_iX_i}{\sum_{i=1}^{n} W_i}$

$\hat{\beta}_0$ and $\hat{\beta}_1$ are the estimates of $\beta_0$ and $\beta_1$ respectively. The expression for their corresponding variances is given by;

$$V(\hat{\beta}_0) = \left[\frac{1}{\sum_{i=1}^{n} W_i} + \frac{(\bar{X}_W)^2}{\sum_{i=1}^{n} W_i(X_i - \bar{X}_W)^2}\right]\delta^2$$

$$V(\hat{\beta}_1) = \frac{\delta^2}{\sum_{i=1}^{n} W_i(X_i - \bar{X}_W)^2}$$

Where, $\delta^2 = \frac{\sum_{i=1}^{n} W_i(Y_i - \bar{Y}_i)^2}{n-2}$
\( \sigma^2 \) is the Mean Squared Error (MSE), \( \hat{Y}_i \) is the predicted value of the dependent variable. Standard error of \( \hat{\beta}_0 \) and \( \hat{\beta}_1 \) is obtained as \( SE(\hat{\beta}_j) = \sqrt{V(\hat{\beta}_j)}, \ j= 0,1 \) and their corresponding adjustment is \( S_j = SE(\hat{\beta}_j)/\sqrt{MSE} \). The 95\% confidence interval (CI) for the fixed effect meta-regression coefficients is given by

\[
\hat{\beta}_j - Z_{(1-\alpha/2)} S_j \leq \beta_j \leq \hat{\beta}_j + Z_{(1-\alpha/2)} S_j.
\]

**Meta-regression: Random effects**

\[
Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i + U_i
\]

This model has two components in its error term namely \( \varepsilon_i \) and \( U_i \), where \( \varepsilon_i \) represents within study variability and \( U_i \) represents between studies variability. Therefore, the variance of \( Y_i \) is \( \sigma^2_i = (\sigma^2 + \hat{\tau}^2) \), where \( \hat{\tau}^2 \) is the estimate of between studies variability. In case of a single covariate, it is estimated as;

\[
\hat{\tau}^2 = \frac{Q - (n - 2)}{F(W, X)} \text{ if } Q > n - 2, \text{ or } 0 \text{ otherwise}
\]

\( Q \) is the heterogeneity statistic, given by \( Q = \sum_{i=1}^{n} W_i(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2 \); \( \hat{\beta}_0 \) and \( \hat{\beta}_1 \) are estimated from (4.8).

\[
F(W, X) = \frac{\sum_{i=1}^{n} W_i - \sum_{i=1}^{n} W_i^2 \sum_{i=1}^{n} W_i X_i - 2 \sum_{i=1}^{n} W_i^2 X_i \sum_{i=1}^{n} W_i X_i + \sum_{i=1}^{n} W_i \sum_{i=1}^{n} W_i^2 X_i^2}{\sum_{i=1}^{n} W_i \sum_{i=1}^{n} W_i X_i^2 - (\sum_{i=1}^{n} W_i X_i)^2}
\]

Then a weighted linear regression is carried out with weights \( W_i^* = 1/\sigma^2_i \) to provide new estimates of \( \beta_0 \) and \( \beta_1 \). The standard errors of the estimated random effects meta-regression coefficients doesn’t require any correction. Thus the 95\% CI is given by \( \hat{\beta}_j - Z_{(1-\alpha/2)} SE_j \leq \beta_j \leq \hat{\beta}_j + Z_{(1-\alpha/2)} SE_j \).
4.2.4.2 Testing the significance of estimated meta-regression coefficients

The statistical significance of estimated coefficients of meta-regression is tested by Z statistic\(^{368}\) of the form:

\[
Z = \frac{\hat{\beta}_j}{SE(\beta_j)}, \quad j = 0, 1
\]  
\[\text{(4.9)}\]

Under the null hypothesis that the coefficient is zero, Z follows normal distribution with mean 0 and variance 1.

The statistical packages estimate the meta-regression coefficients by the method of Restricted Maximum Likelihood (REML), which provides estimates with higher precision.

4.3 Methodology - Third objective

To assess the scope of application of developed methodology in public health evidence summary

Scope of application of developed methodology was assessed by applying it to integrate the findings of studies of 13 public health systematic reviews.\(^{102, 352–363}\) A sensitivity analysis was carried out to determine the robustness of the estimates adjusted for complexity. Advantages and limitations of the developed methodology were determined and also the requirements for further improvement were anticipated.
Key Points

The present investigation involved three objectives;

» First objective was dealt by a qualitative study “Complexity in public health interventions – stakeholders’ perspective: a qualitative analysis”.

» Second objective involved activities namely development of a tool to measure complexity in public health interventions, content validation of the tool, reliability assessment of the tool, identification of probability distribution of the complexity score, classification of studies based on their complexity status and adjustment of the complexity score in meta-analysis using a suitable statistical technique.

» Third objective was accomplished by applying the methodology to integrate the findings of studies of 13 systematic reviews of public health interventions.