



JobNut

Public Health Nutrition Intervention Management



Funded by



Education and Culture

Leonardo da Vinci



PERMISSION FOR USE AND COPYING

Intellectual property rights for this material rests with the Authors.

Permission to use and copy this resource is freely granted for educational purposes. These Units are designed to be delivered intact as complete Units, however there may be situations where use of part of the Unit is warranted in educational contexts. If you are reproducing in part only, please credit:

Hughes R, Black C, Kennedy NP. *Public Health Nutrition Intervention Management: impact and Outcome Evaluation*. JobNut Project, Trinity College Dublin. 2008.

ADDITIONAL COPIES

The complete suite of 18 Public Health Nutrition Workforce Development units developed by the JobNut Project and the related Educator's Guide are available for downloading at the following website:

<http://www.medicine.tcd.ie/nutrition-dietetics/jobnut/>

DISCLAIMER

This workforce development unit has been produced as part of the JobNut Project, supported by the Leonardo Da Vinci Program, Education & Culture, European Commission. The opinions and conclusions expressed in this paper are those of the author(s) and no official endorsement by the funder is intended or should be inferred.

ACKNOWLEDGEMENTS

Funding to support the JobNut Project was provided by the Leonardo Da Vinci Program, Education & Culture, European Commission.

This Unit has been developed by Professor Roger Hughes, Christina Black and Dr Nick Kennedy and of the Unit of Nutrition and Dietetic Studies, School of Medicine, Trinity College Dublin.



TRINITY COLLEGE DUBLIN
COLÁISTE NA TRIONÓIDE, BAILE ÁTHA CLIATH

Intelligence

Action

Evaluation >

Impact and Outcome
Evaluation



Evaluation

Unit 14 - Impact and Outcome Evaluation

Learning Objectives



On completion of this unit, students should be able to:

1. Describe the importance of impact and outcome evaluation in public health nutrition intervention management
2. Explain the difference between impact and outcome evaluation
3. Apply an impact evaluation framework to public health nutrition interventions to measure achievement of intervention objectives
4. Apply an outcome evaluation framework to public health nutrition interventions to measure achievement of intervention goals

Intelligence



Unit Readings

- Macaulay, A., Paradis, G., Potvin, L., Cross, E., Saad-Haddad, C., McComber, A., Desrosiers, S., Kirby, R., Montour, L., Lamping, D., Leduc, N. and Rivard, M. The Kahnawake Schools Diabetes Prevention Project: Intervention, Evaluation, and Baseline Results of a Diabetes Primary Prevention Program with a Native Community in Canada. *Preventive Medicine*. 1997. 26: 779-790.
- Fourney, A. and Bellow, A. 2007. 2005-2006 Impact Evaluation Final Report, *Champions for Change - Network for a Healthy California*. The Research and Evaluation Unit of the Cancer Prevention and Nutrition Section of the California Department of Public Health.
- Valente, T. Statistical Analysis. In *Evaluating Health Promotion Programs*. 2002. Oxford University Press: New York. p.163-180.

Intelligence

Action

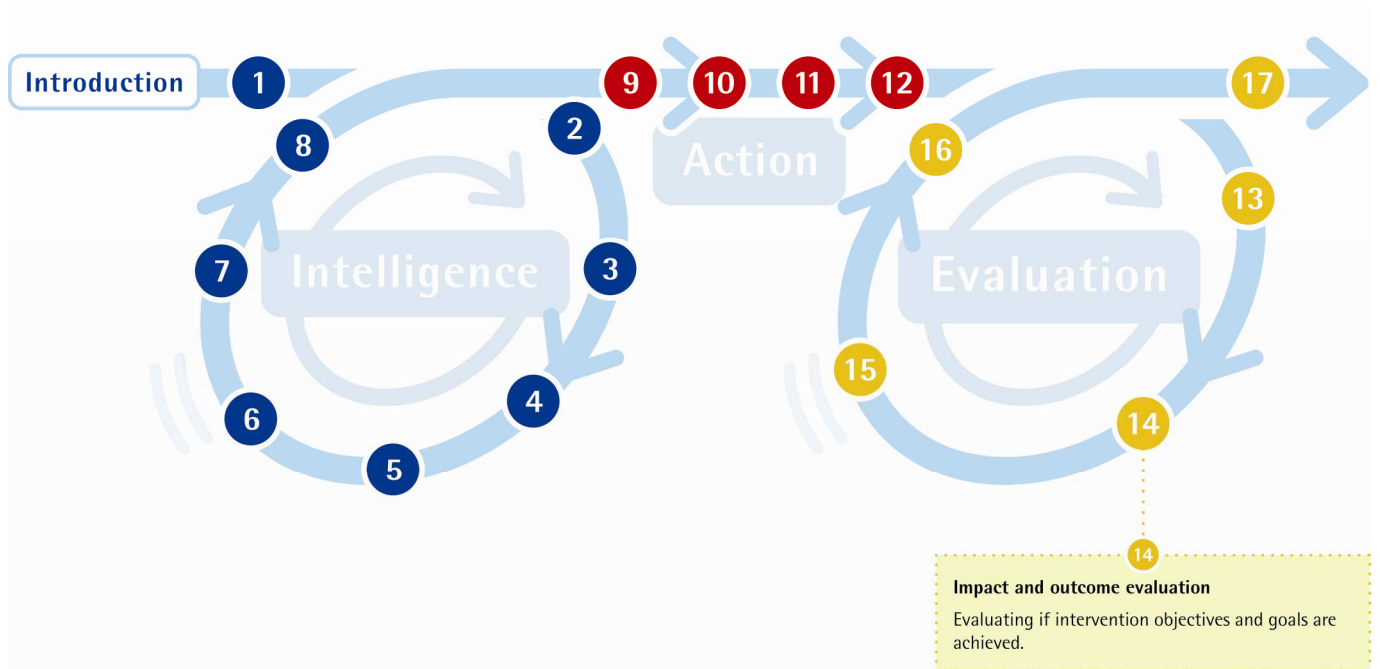
Evaluation >

Impact and Outcome
Evaluation



Contents:

	Page No.
Impact and Outcome Evaluation - What is the Difference?	5
When to Evaluate	6
Key Measures of Impact and Outcome Evaluation	7
Technical Concepts - Reliability and Validity	11
Technical Concepts - Sampling and Data Analysis	12
Evaluation Design	15
Summary of Key Points	17
Additional Resources and Readings	18
References	18



From: Hughes R, Margetts B. *Public Health Nutrition: A Practitioners Handbook*. Blackwell Publishing, 2009.





Impact and Outcome Evaluation

Preamble

Impact and outcome evaluation involve measuring the effects of an intervention, investigating the direction and degree of change. Impact and outcome evaluation test the proposed causal change or logic model of the public health nutrition (PHN) problem by examining change in the identified determinants of the nutrition problem such as behaviour, attitudes, health status, environmental and societal factors etc.

Impact evaluation is concerned with the assessment of the immediate effects of the intervention and usually corresponds with the measurement of intervention objectives.

Outcome evaluation is concerned with measuring the longer-term effects of the intervention and this usually corresponds to the intervention goal. The key difference between impact and outcome evaluation is not what is being measured but rather is defined by the sequence of measurement. Whether a factor is measured in impact or outcome evaluation is entirely dependent upon the causal chain of events postulated and what aspects of this causal chain the intervention is aiming to address.

In this unit the process of completing impact and outcome evaluation is outlined, Various tools to assist with measuring evaluation indicators, design types and technical considerations required for evaluation are explained.

Impact and Outcome Evaluation - What is the Difference

Impact and outcome evaluation both involve the assessment of intervention effects but at different levels. Impact evaluation is concerned with the assessment of the immediate effects of the intervention and usually corresponds with the measurement of the intervention objective. Outcome evaluation is concerned with measuring the longer-term effects of the intervention and this usually corresponds to the intervention goal (1). For example, a community-based nutrition intervention may be attempting to increase knowledge, awareness, and availability of fruit and vegetables with the goal of increasing fruit and vegetable intake by an average of one serving across the community. Impact evaluation would assess changes in knowledge, awareness, and availability of fruit and vegetables while outcome evaluation would measure variation in fruit and vegetable intake behaviour.

Impact and outcome evaluation test the logic model or causal chain of events that has been postulated by the intervention. Such as changing knowledge, awareness and availability will lead to a change in dietary behaviour.

As covered previously in Unit 4 intervention logic models are derived from analysis of the population nutrition problem, and its determinants, the factors that the intervention set out to address.

The key difference between impact and outcome evaluation is not what is being measured but is rather defined by the sequence of measurement. Whether a factor is measured in impact or outcome evaluation is entirely dependent upon the causal chain of events postulated and what aspects of this causal chain the intervention is aiming to address (1).

Intelligence

Action

Evaluation >

Impact and Outcome
Evaluation



A factor that is assessed in outcome evaluation in one intervention may be assessed as part of impact evaluation in another. For example, a public health nutritionist working in community health may be trying to affect change in knowledge and availability of fruit and vegetables (impact) with the goal of increasing average intake by one serving (outcome). While at the national or pan-European level increasing fruit and vegetable intake (impact) may be an objective towards achieving the goal of reducing the prevalence of obesity.

It is important however not to generalise thinking that intervention goals and outcome indicators have to be stated in terms of health status, as is sometimes suggested (1). PHN interventions at all levels should follow through the factors in the causal chain they are trying to address by measuring both the immediate effects (impact) and the subsequent effects (outcome) no matter what factors are being measured at each point. If interventions were restricted to the assessment of immediate effects only, intervention success would become narrowly and prematurely defined. Thorough impact and outcome evaluation is needed to determine when the causal theory does and does not hold (1).

With some interventions or intervention strategies impact and outcome evaluation will be focused on what the project management committee or community consider success to be and how it will be brought about. This type of success measure is particularly appropriate in capacity building interventions and strategies, and could include changes in perceived power and community self-confidence, concern for local issues or the formation of action groups. Methods for measuring community capacity gains are addressed in more detail in Unit 15 Evaluating Capacity Gains.

When to Evaluate?

Predicting when the intervention effect/s will take place and the timing of impact and outcome evaluation is very important because measuring effects too early or too late will deliver false findings about the intervention's effectiveness (1). There are several possible effects an intervention can have over time:

- **Ideal effect** - the intervention has an immediate improvement in the nutrition factor of interest which is sustained over time.
- **Sleeper effect** - the intervention impact is not detected until sometime after implementation such that the effect would be missed if evaluation is undertaken immediately or after only a short while.
- **Backsliding effect** - the effect of the intervention is immediate but only short lasting. The effect would be missed if measured only sometime after implementation.
- **Trigger effect** - an intervention which triggers or brings forward a behaviour or event that would have happened anyway. Effects are seen immediately after implementation, then drops below baseline level before returning to normal or baseline level.
- **Historical effect** - cases where a health behaviour or target factor is gradually improving across time such that the evaluation captures this effect and wrongly attributes it to the intervention. Highlights the need for external or secular trends to be distinguished from intervention effects.
- **Backlash effect** - occurs when premature cessation of an intervention demoralises or disillusions participants leading to levels of behaviours or problems that are worse than baseline. Immediate evaluation only would miss the subsequent negative effects (1).

Intelligence

Action

Evaluation >

Impact and Outcome
Evaluation



In order to be informed about which sort of effect could apply to your intervention it is important to consult the literature for previous or similar interventions with the same target population, and seriously consider the potential and likely effect the intervention will have. Is the intervention likely to have an immediate effect that could taper off or is the intervention likely to take some time to show the intended effect? In cases where predicting the likely effect is total guess work a pilot study should be implemented with multiple measures overtime will elicit a predicted effect and assist with implementing the main intervention (1).

Exercise 1.



Thinking about the possible timing of intervention effects explained above consider your selected scenario and write a brief description of what effect you predict your intervention or various intervention strategies will have. You will need to refer to the literature for guidance or outline a pilot study methodology that will assist your prediction.

Workshop/tutorial option:

Complete the exercise in small groups followed by a whole-class debriefing

CPD option:

Conduct the above exercise in the context of your current work role and an identified nutrition problem in the community or population you are working with.

Key Measures of Impact and Outcome Evaluation

There are some common measures used in impact and outcome evaluation in PHN interventions that are outlined below. A mixture of qualitative and quantitative methods is used in these evaluation measures. Qualitative methods interpret the meaning of the intervention for participants, staff, key stakeholders, and also those not reached by the intervention (1). Qualitative methods are largely unstructured and observational. Quantitative methods systematically measures the size of the intended effects of the intervention using standardised measuring instruments, scoring these measures and then undertaking statistical analysis on these scores (2). Having SMART (see Unit 9) objectives and goals makes analysis of the quantitative data considerably easier as the level of intended effects is clearly specified. Qualitative approaches ask 'why' and quantitative approaches examine 'how much'.

In the evaluation and intelligence phases both types of approaches are used to interpret what is happening. The extent that each method is used is dependent upon the intervention strategies, the target group and the size of the intervention. It is most effective however to use a combination of qualitative and quantitative approaches.

Intelligence

Action

Evaluation >

Impact and Outcome
Evaluation



Knowledge

Measuring knowledge involves assessing what people know, what people recognise, what they are aware of, what they understand and what people have learned. Measuring knowledge is commonly broken into measuring awareness or recognition of an intervention or intervention message, and measuring what the target audiences have understood or have learnt about an issue or subject area.

Awareness of an intervention or key message is commonly used for measuring social marketing and communication campaigns. This form of measurement can involve presenting representatives of the target group with a list of different campaigns/messages (a combination of real and fictitious) that includes the intervention and asking them about their awareness of each item on the list. Measuring awareness using a variety of prompts avoids the ‘demand characteristic’ of asking a more direct question about the intervention, such as ‘are you aware of a recent healthy eating campaign’ to which most people will respond ‘yes’ (1). However using an open ‘unprompted’ question like ‘can you name any healthy eating campaigns you have seen in the past six months’ can be used before providing a list of prompts.

Assessing knowledge about an intervention or subject matter can be undertaken several ways, three of which are described below. It is important to consider when measuring knowledge that people, unless agreeing to be explicitly involved in a learning situation such as a training course, tend not to respond positively when asked ‘what did you learn?’ Most people tend to think they do not need to be taught anything.

Three ways to measure knowledge of a subject matter:

- Recall - assessed by asking ‘what did the campaign/ intervention encourage people to do?’
- Learning - participants are asked to mark statements about a particular topic as true or false or select the correct response from a series of alternatives (multiple choice format).
- Change in knowledge - before and after questionnaires using a similar format as above. Take care with questionnaires which ask people to rate the knowledge on a subject as awareness of complexity of a topic can cause people to become more modest in their self-assessment post intervention (1).

Attitudes and self-efficacy

Measuring attitude and self-efficacy involves assessing how people feel about the intervention or topic matter, or their ability to participate in intervention activities and commonly involves qualitative methods which encourage more freedom in expression (1). Possible methods of exploring attitudes can include showing short films, roll-plays or picture/verbal stories of scenarios depicting the topic of interest. For example showing a family who is eating their evening meal of take-away with soft-drink in front of the television and asking ‘what do you think about this family’s evening meal habits and the foods they are consuming?’ Responses can be collected through individual interviews, a focus group or a questionnaire similar to that shown in Table 1. Self-efficacy can be measured in a similar manner by enquiring about how the individual feels about their ability to undertake and continue intervention activities.

Intelligence

Action

Evaluation >

Impact and Outcome
Evaluation



Table 1. Questionnaire for measuring attitudes

<i>Please circle the response that most closely matches how you feel about each statement</i>	
Eating the evening meal in front of the television promotes healthy eating habits	Strongly agree - agree - don't know - disagree - strongly disagree
Eating take-away foods five times a week is healthy	Strongly agree - agree - don't know - disagree - strongly disagree
It is healthy for families to eat the evening meal together at the dining table	Strongly agree - agree - don't know - disagree - strongly disagree

Adapted from: (1)

Behaviour

The purpose of most PHN interventions is to bring about change in people's dietary and activity behaviour, to eat more nutritious foods or exercise more. Measuring behaviour can be achieved through self-report however this method is generally not accurate because of the social desirability attached to healthy eating and engaging in regular physical activity. Asking people to record their behaviour overtime in a food and exercise diary can minimise inaccuracy however can also influence behaviour, making the reporting process more a part of the intervention than of the evaluation (1).

Observation is another method that can be used to measure behaviour. Although sometimes considered inconvenient observation may be the most effective method of measuring behaviour (1). Various forms of observation may include watching the food choices employees or students make at a factory or school canteen, asking participants to photograph their dinner plates on a specially designed placemat before eating, or calculating product consumption from sales data.

Health status

When selecting an appropriate health status measure for impact or outcome evaluation it is important to first revisit the intended effect of the intervention and then to ensure the selected measure includes the necessary dimensions and is suited to the target group. In PHN interventions health status can be measured using biochemical or anthropometric indicators for diet-related conditions such as obesity, glucose intolerance, diabetes and cardiovascular disease, or measures of physical fitness.

Social support

PHN interventions may attempt to increase the quality and quantity of social support given to a particular target population such as young parents, newly arrived immigrants, or people living in disadvantaged communities. There are a variety of self-completed questionnaires and interview schedules available to measure social support. Selecting which measure is most appropriate depends upon which aspect of social support the intervention is emphasising such as the extent to which the social support provides information, practical assistance, self-efficacy, or meets the individual's expectations.

Intelligence

Action

Evaluation >

Impact and Outcome Evaluation



Simple measures can also be used for example, the number of young mothers who can provide the names of each other's partners or children or if they have visited each others homes since commencing the intervention (1).

Environmental support

Many comprehensive PHN interventions aim to bring about change at an environmental level as well as changing social and behavioural factors. Measuring environmental support considers change in the physical environment, policies, legislation and workforce support. For example, an intervention targeting physical activity in the workplace may measure environmental support by auditing the work environment considering the availability of secure bike racks, shower, locker or gym facilities, accessibility of stairwells, and changes in organisational policy about active commuting, flexible work hours or salary packaging. Environmental audit tools for different surroundings such as schools, workplaces and communities are becoming more readily available.

Intelligence

Reading



Macaulay, A., Paradis, G., Potvin, L., Cross, E., Saad-Haddad, C., McComber, A., Desrosiers, S., Kirby, R., Montour, L., Lamping, D., Leduc, N. and Rivard, M. The Kahnawake Schools Diabetes Prevention Project: Intervention, Evaluation, and Baseline Results of a Diabetes Primary Prevention Program with a Native Community in Canada. *Preventive Medicine*. 1997. 26: 779-790.

Fourney, A. and Bellow, A. 2007. 2005-2006 Impact Evaluation Final Report, *Champions for Change - Network for a Healthy California*. The Research and Evaluation Unit of the Cancer Prevention and Nutrition Section of the California Department of Public Health.

Exercise 2.



After reading the articles by Macaulay et al (1997) and Fourney and Bellow (2007) and considering the impact and outcome evaluation strategies and tools used to assess the interventions objectives and goals. Review your PHN scenario objectives and goals and impact and outcome indicators (developed in Units 9 and 11 respectively). Draft or search for a tool to measure your outcome indicator and an impact indicator.

Workshop/tutorial option:

Complete the exercise in small groups followed by a whole-class debriefing

CPD option:

Conduct the above exercise in the context of your current work role and an identified nutrition problem in the community or population you are working with.

Intelligence

Action

Evaluation >

Impact and Outcome
Evaluation



Technical Concepts - Reliability and Validity

While it is not possible to detail the process of constructing the various evaluation tools outlined above (questionnaires and scales, interviews and focus groups) there are two important technical concepts to consider when developing evaluation tools to assess changes in knowledge, attitudes, behaviours and environmental or social factors; reliability and validity.

Reliability is the stability of a measure. A reliable tool measures the same things each time the measure is used and for each person it is used with. The most common method used to test and develop reliability is to repeat administration of the measurement on the same subject using the same administration procedures within a short period of time to ensure this 'test-retest' procedure elicits the same results (2). If the same participant answers a set of questions in relation to a specific behaviour or nutrition problem in the same way within a relatively short period of time the measurement tool can be considered reliable. Testing for reliability is also important for observable indicators, such as observing or rating the physical activity characteristics of a community. Reliability of an observation tool can be determined by the level of agreement between two observers of the same phenomenon, known as 'inter-rater reliability'. If the observations differed too greatly, the measurement tool would be unreliable and strict guidelines would need to be developed for the manner in which data is collected or the number of observers can be limited to minimise variations in observation (2).

Validity is the truth of a measure. A valid tool is a tool that measures what it intended to measure. A common approach to assess validity is the use of biochemical or physiological tests, where these tests are considered 'true' measures of the factors of interest. For example, the measurement of blood lipids to validate aspects of self-reported diet. Although desirable, such measures are generally difficult and expensive, not always practical to use and are only available to support measures of behaviour.

Measurement of attitudes and beliefs, and psychological concepts such as self-efficacy (very important for increasing physical activity) are not able to be objectively assessed however simple procedure of face validity and content validity can be employed to check validity. **Face validity** involves experts in the field agree the measure is a useful way of assessing the phenomenon in question. **Content validity** ensures all of the areas of the phenomenon in question are covered by measurement items (2). For example, a valid questionnaire would present options in a balanced manner and that allows respondents to answer across the full range of potential responses. By asking respondents the degree to which they agree or disagree with a statement the respondent is able to express their answer in different degrees, and the evaluator is able to analyse the difference in response in a more sophisticated, systematic manner. This form of question design presents potential responses in the form of a ranked scale ('strongly agree', 'agree', 'neither nor disagree', 'disagree', 'strongly disagree') and is called a *Likert scale*.

Valid measures of concepts such as social capital or capacity building are more challenging to develop because concepts are factors in a causal chain that are not directly observable. Being not directly observable turning concepts into actual measures is a technical process derived from psychology and other social sciences. The aim is to turn concepts into variables that can reliably and validly show variation between subjects and variation as a consequence of intervention (2). For example, there is no direct measure for observing community capacity however this concept can be described and relevant questions generated, administered to the target group and psychometric statistical techniques used to describe how well the questions relate to the construct or how well the construct

Intelligence

Action

Evaluation >

Impact and Outcome
Evaluation



exists. This process is known as **construct validity**. Techniques for measuring capacity gains are further explained in Unit 15.

It is important to note that a reliable tool is not necessarily a valid tool as one can be measuring the wrong thing but be doing so consistently. It is also worth researching the literature for published instruments or measures before commencing to design a new measure. Using available measures will enable a comparison with already published data and save considerable time and resources.

Exercise 3.



Considering your results to Exercise 2, explain how you would test the reliability or validity of your selected evaluation tools. If the tools you have selected are already reliable and valid provide details about the author, previous application tool and process the tool underwent to become reliable and valid.

Workshop/tutorial option:

Complete the exercise in small groups followed by a whole-class debriefing

CPD option:

Conduct the above exercise in the context of your current work role and an identified nutrition problem in the community or population you are working with.

Technical Concepts - Sampling and Data Analysis

Impact and outcome evaluation designs should consider potential sources of bias, sampling methods and methods of data analysis, each of which are outlined below.

Sampling bias

Bias is where something differs systematically from the true situation and influences the evaluation conclusions. Sampling bias concerns the characteristics of intervention participants, reasons for their participation and the duration of their participation. How participants are recruited to participate in the intervention and whether or not they are representative of the whole target population has an important impact on the evaluation findings. For example, those who volunteer to participate in an intervention or evaluation may have greater motivation or health literacy than those who do not volunteer and can exaggerate the intervention results by producing better outcomes than if a more representative population participated in the intervention. This type of bias is particularly important to address in intervention trying to have the greatest impact amongst more marginalised or disadvantaged populations.

Sampling bias can also occur as a consequence of non-response, when a person appropriate to participate in the intervention refuses to do so. This form of bias also produces different effects than would be observed if the program were delivered to a fully representative population. For example, if an intervention was targeting 200 maternal and child health nurses in a particular region however only

Intelligence

Action

Evaluation >

Impact and Outcome
Evaluation



20 actually participated, the intervention would elicit a non-response bias. The consequence of both these forms of sampling bias is that the results cannot be generalised and extrapolated to the target population as a whole.

Another source of sampling bias concerns subject retention. This type of bias can occur if participants who drop out of the intervention or participate in the baseline data collection only are different to those who complete the intervention and/or the follow-up data collection. The extent of drop out or loss to follow-up can influence the usefulness and generalisability of the results.

It is important to do all that is practically possible to obtain a representative sample and maintain participation rate throughout the intervention to avoid sampling bias. Strategies to retain participants can include using easy-to-complete questionnaires, on-going communication or incentives. When analysing impact and outcome evaluation results any sampling bias should be described and possible influence on results outlined.

Sampling methods

Sampling methods for impact and outcome evaluation tend to differ according to the size of the PHN intervention. In small scale interventions it may be possible to measure the intervention effect in all participants. Most PHN interventions however, rely upon a subset of individuals from the population to assess the impact and outcome of the intervention. A random sample is considered the best method for evaluating intervention effects in large population groups because the effects of the intervention can be considered as applicable to the entire target population. Taking a random sample requires a list of the whole target population to be available and that it is possible to select subjects at random. Examples of population lists include census lists, employee payroll lists, school lists of enrolled students or telephone listings. It may not always be possible or practical to achieve a true random sample. Table 2 illustrates different types of sampling that can be used for impact and outcome evaluation of PHN interventions.

Intelligence

Action

Evaluation >

Impact and Outcome
Evaluation



Table 2. Sampling and recruitment methods for evaluation of large and small PHN interventions

	Recruitment of participants - small PHN intervention ¹	Sampling of participants - large PHN intervention ²
Best sampling method (generalisable)	Random sampling from target population of newly arrived immigrants	Random sample of children aged 5-12 is measured - sampling of the at-risk whole population allows every individual an equal chance of selection
↓	Sampling from a defined database of older people	Other variants: random sampling with increased samples (oversampling) of specific groups of interest, particularly those attending schools in more disadvantaged areas; universal sampling where everyone is surveyed because the target population is small
	Sampling from numerous community groups and clinical settings - even if non-random may be adequately diverse to be generalisable	
	Less useful sampling method (less reliable)	Snowball samples where hard-to-reach groups are found through social networks - can produce reasonable samples
	Volunteer immigrants recruited through newspaper/newsletter advertisements	

Adapted from: (2)

Statistical analysis

It is essential to consider the use of statistical methods to analyse and make sense of the quantitative data collected from impact and outcome evaluation. Statistical analysis allows evaluation data to be interpreted and produces useful information about the success or otherwise of an intervention. Statistical methods should be taken into account during evaluation planning to determine the sample size of the evaluation and which statistical tests to apply. Some key statistical considerations are outlined below:

- **Statistical significance** - the probability of the observed result occurring by chance. Often described as *p* values of <0.05 or <0.01 indicating there is a 1 in 20 or 1 in 100 possibility of an observed outcome by chance respectively.
- **Confidence intervals** - describe how likely the true population results are to be outside the range described by the confidence limits.

Different statistical tests are required for continuous data such as daily vegetable intake (*t*-tests and Pearson coefficients) compared to category data such as vegetable intake improved/ did not improve (chi-squared statistics and odds ratios). The reading below expands further on statistical analysis for evaluation.

¹ For example, trial of a community kitchen intervention to increase health eating knowledge and skills with newly arrived immigrants

² For example, a national school program to increase reduced-fat dairy intake and physical activity levels in children aged 5-12 years





Practice Note



The ‘best approach’ to PHN intervention evaluation varies according to the context and setting of the intervention, the resources and time available, and the requirements for evidence of intervention effectiveness.

A general rule for costing evaluation in funding submission and intervention plans is 20% of the total intervention budget. However this level of funding is not always available and when the evaluation budget is limited evaluation designs tend to rely more heavily on the use of information describing the process of implementation (process evaluation), and qualitative interpretation of observed changes in the target group.

Intelligence



Reading

Valente, T. Statistical Analysis. In *Evaluating Health Promotion Programs*. 2002. Oxford University Press: New York. p.163-180.

Exercise 3.



After reading the chapter by Valente (2002), consider your results to Exercise 2 and outline the methodology you will undertake to statistically analyse your impact and outcome evaluation data.

Workshop/tutorial option:

Complete the exercise in small groups followed by a whole-class debriefing

CPD option:

Conduct the above exercise in the context of your current work role and an identified nutrition problem in the community or population you are working with.

Evaluation Design

While randomised control trials are considered the gold standard in health science because they are perceived as reliable and valid their use in health promotion and public health interventions is limited (3). There are a number of possible impact and outcome evaluation designs which are outlined below in Table 3.

Intelligence

Action

Evaluation >

Impact and Outcome
Evaluation



Table 3. Possible evaluation designs for PHN interventions

Design type	Description	Considerations
Single group, post test only	A single measurement at the completion of the intervention is taken on the participants only	Unsure whether the intervention actually had an effect and that any effect observed was actually caused by the intervention
Single group, pre- and post-test	Two measurements are taken, one before and one after the intervention, on the participants	Pre-post design enables change to be observed but cannot be sure other factors caused or contributed to the effect
Non-equivalent control group, pre- and post-test	Two measurements are taken, one before and one after on two groups, an intervention and a control group. 'Non equivalent' means the two groups are not exactly matched in characteristics and may be from another region, time period etc	Pre-post design enables change to be observed however because the groups are not matched it may be group differences rather than the intervention which cause the effect. It is possible to statistically control for known differences.
Single group, time series	Multiple measures over a period of time on participants only	Able to observe natural changes occurring in the participant group, and the size and direction of these and changes before implementing the intervention and then observe the effects. Important to consider possible external influences during the time period
Non-equivalent, time series	Multiple measures over a period of time on two groups, an intervention and a 'non equivalent' control group	Stronger evidence than the above design because more definitely rules out external factors and help to cancel secular trends that can be missed with pre-post designs.
Randomised control trial	Two measurements are taken, one before and one after on two groups, an intervention and an equivalent control group.	Overcomes all issues of the above designs.

Adapted from (1, 3)

Selecting an evaluation design for an intervention is dependent upon the context, particularly the financial and staffing resources and the baseline data available. Although some designs are considered more scientifically sound than others because they elicit stronger evidence about intervention effects, some designs are easier to execute, hence design selection is about finding the balance suitable for the context.

Intelligence

Action

Evaluation >

Impact and Outcome
Evaluation



Assessment

Considering your selected scenario and using your responses to Exercises 2, 3 and 4 complete the impact and outcome evaluation section of the intervention management template.



CPD option:

Conduct the above exercise in the context of your current work role and the community or population you are working with.

Key Points



- Impact and outcome evaluation involve measuring the effects of an intervention, investigating the direction and degree of change and testing the proposed causal change or logic model of the population nutrition problem.
- Impact evaluation is concerned with the assessment of the immediate effects of the intervention and usually corresponds with the measurement of the intervention objective. Outcome evaluation is concerned with measuring the longer-term effects of the intervention and this usually corresponds to the intervention goal.
- Predicting when the intervention effect/s will take place and the timing of impact and outcome evaluation is very important because measuring effects too early or too late will deliver false findings about the intervention's effectiveness.
- There are some common measures used in impact and outcome evaluation in PHN interventions including knowledge, attitude and self-efficacy, behaviour, health status, social support and environmental support. A mixture of qualitative and quantitative methods is used in these evaluation measures, and where possible these measures should be reliable and valid.

Intelligence

Action

Evaluation >

Impact and Outcome
Evaluation



Additional Resources and Readings

Impact/Outcome evaluation methods

- Hawe, P., Degeling, D. and Hall, J. *Evaluating health promotion: a health worker's guide*. 1990. MacLennan and Petty: Sydney.
- Valente, T. *Evaluating Health Promotion Programs*. 2002. Oxford University Press: New York. p.87-162.
- Thorogood, M. and Coombes, Y. *Evaluating health promotion: practice and methods*. 2004. Oxford University Press: New York.

References

1. Hawe, P., Degeling, D. and Hall, J. *Evaluating health promotion: a health worker's guide*. 1990. MacLennan and Petty: Sydney.
2. Nutbeam, D. and Bauman, A. *Evaluation in a nutshell: a practical guide to the evaluation of health promotion programs*. 2006. McGraw-Hill: Sydney.
3. Britton, A. and Thorogood, M. In Thorogood, M. and Coombes, Y. *Evaluating health promotion: practice and methods*. 2004. Oxford University Press: New York.

Intelligence

Action

Evaluation >

Impact and Outcome
Evaluation