Evaluation Design

The evaluation design describes the set of procedures and tasks that need to be carried out to systematically examine the effects of a programme (Nutbeam and Bauman, 2007).

The evaluation design of a health promotion programme can be broadly grouped into 3 main types, based on the strength of evidence they provide for intervention effectiveness.

- Experimental
- Quasi-experimental
- Non-experimental

This section aims to provide a brief overview of each evaluation design and highlight the advantages and disadvantages of each approach in relation to the evaluation of exercise referral schemes.

Experimental Evaluation Design

Experimental designs are regarded as the most rigorous and scientific approach to evaluation of effectiveness and the Randomised Controlled Trial (RCT) is often considered as the 'gold standard'.

Randomised Controlled Trial (RCT)

An RCT involves assigning eligible participants at random to either an 'intervention' or a 'non intervention' group. As the assignment of groups is by chance, it is assumed that the groups are similar on all other characteristics that might effect the outcome measures of interest. RCT's are designed to have a very high level of 'internal validity' but because of the way they are conducted they may have less 'external validity' or application in the real world. Because an RCT uses random allocation it is assumed that this will minimise any differences between groups at the start of a programme and allow any observed changes in the intervention group to be attributed to the intervention and not due to 'chance effects' that have nothing to do with the programme. In other words an RCT improves the confidence that the observed changes were *caused* by the programme.

Clustered Randomised Control Trial

Cluster randomisation involves participants being allocated to either the intervention or a control condition, as a group, rather than individually. In the context of an exercise referral scheme, patients could be cluster randomised by practice. Participants referred from one practice receive the intervention and patients from another practice form the non-intervention control group.

Cluster randomisation may be more practical than randomising on the individual level, and may also reduce the risk of contamination.

Quasi-experimental Evaluation Design

Quasi-experimental designs are commonly used in the evaluation of programmes when random assignment is not possible or practical. Like experimental designs, quasi-experimental designs involve comparing the changes between one group that receive the programme and a no intervention control group. The decision of who receives the programme and who doesn't is not random and is usually determined by either systematic allocation or convenience. Although it is desirable for the comparison group to be as similar as possible to the intervention group on factors which could affect the selected outcomes, for instance age and gender, this may not always happen.

In an exercise referral context, like in the cluster randomised design, participants referred from one practice could receive the intervention and patients from another practice form the non-intervention control group, however in a quasi-experimental design the choice of which practice is allocated to the intervention and to the control is not random.

Non-experimental Evaluation Design

Although non-experimental designs are used to evaluate health promotion programmes, including exercise referral schemes, these designs have the lowest level of scientific quality and the least confidence that the changes were caused by the intervention. As a result, these types of studies are often excluded from systematic reviews of evidence. This was the case with the NICE review of exercise referral schemes (NICE, 2006) and similar reviews of interventions conducted in primary care in the USA and elsewhere (U.S. Preventive Services Task Force).

A non-experimental design does not involve a control group and therefore provides limited evidence that the observed changes were due to the intervention or programme and not due to other influences.

Which Evaluation Design Should I Use?

This table illustrates some of the advantages and disadvantages of different evaluation designs in relation to exercise referral schemes.

Evaluation Design	Advantages	Disadvantages
Experimental	Provides the strongest level of confidence that observed outcomes are a result of the intervention	Requires a control group who would not receive the exercise referral intervention Delaying or denying access to a programme requires ethical approval Likely to require specialist expertise to set up and oversee evaluation design, specifically random allocation, maintain quality control of intervention, and avoid contamination of the control group Difficult to control all the variables which may influence programme outcomes Requires a larger budget
Quasi-experimental	The context and way in which the intervention or programme is delivered is often more natural or more similar to 'usual practice' than in an RCT, meaning the results may be more generalisable to other schemes/settings	The comparison group may differ considerably from the intervention group at the outset, making the interpretation of results problematic
Non-experimental	Lends itself to the evaluation of the 'real-life' situation More feasible to implement in the context of exercise referral schemes	Does not provide compelling evidence that the intervention *caused* the observed changes Findings may not be generalisable to other exercise referral schemes