Increasing the use of experimental methods in nursing education research

ABSTRACT

Aim: To consider how more use could be made of experimental research in nursing education.

Background: Much use has been made in nursing educational research of before and after, within-subjects research. While this has its place and has been a valuable design for testing educational interventions, there has been a distinct lack of more sophisticated experimental designs.

Design: Discussion paper. Our approach was to consider the use of a range of experimental designs in nursing education research. The designs included within-subjects designs, between-subjects designs and new approaches to experimental research with human participants such as pragmatic designs, non-inferiority designs and the framework offered by complex interventions.

Conclusions: Within-subjects designs have dominated experimental research in nursing education. While suitable for preliminary studies, these studies should find less acceptance in nursing education journals, and we should see them replaced by more rigorous designs based on between-subjects designs. These do not have to be strictly randomised controlled trials and there are many reasons why these are hard to implement in nursing education research. However, a range of alternatives are available.

Keywords:

Nursing education

Experiments

Complex interventions

Non-inferiority trials

Pragmatic trials

Introduction

As evidenced from the pages of nursing education journals and other nursing journals which publish educational research a range of methods is used to investigate nursing education (Long & Johnson, 2002; Schneider et al., 2013; Morton, 2017) These include both qualitative and quantitative designs and within the quantitative designs a range of methods from correlational research through to experimental research. While there has been an increase over the years in the use of experimental designs in clinical nursing research (Baldi et al., 2014), there is a dearth of rigorous experimental designs being used in nursing education, particularly randomised controlled trials (RCTs; MRC 1948). For example, a cursory search shows that most issues of *Journal of Advanced Nursing* (*JAN*) will contain several RCTs and many issues of *Journal of Clinical Nursing* (*JCN*) likewise. In fact, each of these journals has published hundreds if RCTs as evident if these are specifically searched for in each journal. However, in the two leading nursing educational journals *Nurse Education Today* (*NET*) and *Nurse Education in Practice* (*NEP*), of which the present authors are all editors, experimental studies are rare and RCTs, as a sub-set of experimental studies, are even more rare. In the past 5 years 6.8% of the content of *NET* has been represented by experimental work (58 studies) and RCTs represent 1.8% of the content. The respective figures for NEP are 2.2% (23 studies) and only 0.8% of the content was RCTs. In the corresponding period *JAN* will have published several hundred. Clearly, it could be expected that a generic journal such as *JAN* and a clinical journal such as *JCN* would publish more RCTs. However, we wish to explore the paucity of RCT type studies in nursing education and to compare designs.

It is not entirely clear why there is a paucity of the more rigorous designs in nursing education research. Within-subjects studies are easier to conduct than between-subjects studies. Within-subjects studies can conveniently be conducted in a single site and require smaller sample sizes than between-subjects studies. As such they are cheaper to conduct which is possibly a driver as the funding allocated to nursing research is orders of magnitude less than that allocated to, for example, medical research; approximately $20 million as opposed to the approximately £50 billion (NINR, 2022). In our experience very little of that is allocated to nursing education research. It is also our experience that some nurses are express anxiety towards statistics generally (Hagen et al., 2013) and develop a fear of the advanced statistical methods required to design and analyse experimental studies.

The term ‘experiment’ is general and simply describes and investigation designed to test an hypothesis. Designs vary in their ability to distinguish between cause and effect and extent to which they are externally valid. External validity refers to the extent to which the outcome of a study may be generalised to other situations (Steckler & McLeroy, 2008). The design which best distinguishes cause and effect is the true experiment where a group receiving an intervention is compared with one that does not receive the intervention (a control group). Ideally the allocation to the two groups is random and participants and data collectors are blind to the allocation. However, all aspects of a true experiment described above are not always possible and any compromise in control, randomisation or blinding creates a quasi-experiment (Campbell, Stanley & Gage, 1963). The relationship between rigour and external validity is reciprocal and it must be acknowledged that true experiments are low in external validity.

Prior to embarking on our consideration of a range of experimental methods (Table 1), we wish to emphasise that for all the designs we consider, the need for clearly defined outcome variables and rigorous measurement is essential.

Within-subjects designs

Within-subjects designs (or repeated measures) are attractive to educational researchers for several reasons. The main reason is convenience in the sense that a group of students who can be tested before and after an intervention is relatively easy to find. These designs are also statistically powerful as they do not suffer from individual variation; participants serve as their own controls. But they also suffer from several methodological problems not the least of which is that they are low in external validity.

They also suffer from and are affected disproportionately compared with between-subjects research by participant fatigue and participants may drop out between the two waves of the study. It is also hard to account for the maturation effect of changes unrelated to the intervention and not under the control by the researchers, but which may, nevertheless, influence the outcome of the study. Moreover, the pre-post measures could be also affected by a recall bias related to the participants’ previous knowledge of the scale and questionnaires adopted. A longitudinal design employing multiple measures on the same participants should carefully consider these biases. While a longer timespan from measure to measure could mitigate the recall bias, it affects the potential participants’

dropout and missing data. Some statistical approaches, as the differential dropout and bias tests, could mitigate the effect of data missing in longitudinal designs (Bell et al., 2013), however, a rigorous study design is preferred to mitigate biases instead of controlling them in the analytic phase. Clearly, in within-subjects research, the participants cannot be blinded to the fact that they are having the intervention, and, without a comparison group, the effect of the intervention cannot be isolated.

Nevertheless, within-subjects rather than between-subjects research have a valuable role in developing educational interventions and testing the feasibility of interventions but are inadequate for truly testing the efficacy of interventions.

Between-subjects designs

Researchers using between-subjects designs can overcome many of the problems of within-subjects research. Principally, they have the potential to be more externally valid and this can be achieved by blinding, control and randomisation.

Blinding

Blinding in between-subjects designs can take several forms. The gold standard for clinical trials is double blinding whereby neither the participant nor the data collector is aware that, respectively, they have undergone the intervention or who has had the intervention. Clearly, as with within-subjects designs, blinding of participants in educational research is more challenging. Nevertheless, it should be eminently possible in well-designed studies to achieve single blinding whereby, at least, the data collectors are unaware of who has undergone the intervention (Carbogim et al., 2018)

Control

Control in between-subjects research involves having a group who do not undergo the intervention, often referred to as a control group. Ideally, a control group undergoes all aspects to the study except the intervention. However, in educational research there are several issues that can make proper control problematic. One issue is simply the difficulty of finding a suitable control group. In large classes of students, it should be possible to allocate students to an intervention and a control group in theory. In practice this may be harder and even if students are so allocated then practical difficulties may be encountered in applying the intervention to one group and withholding it from another. This may require more than one classroom and physically dividing the group. But even if this is achieved the phenomenon of contamination of the control group by the intervention group may take place

(Keogh-Brown et al., 2007). While it can be difficult to achieve, some small classroom based studies have used control groups for specific interventions such as in the study designed by Padiha et al. (2019) about clinical virtual simulation or Tamaki et al. (2019) for an end-of-life educational intervention in nursing education. However, it is hard to prevent the students communicating outside class and exchanging information. In educational research, cluster randomisation is more suitable, meaning that the control group can represent participants from different organization. In such cases, baseline characteristics of the participants need to be comparably similar with experimental group.

Furthermore, in common with clinical research, the issue of what constitutes a good control group arises. It would be unethical and not feasible to withdraw educational input to one group while the other received a novel intervention. It is more likely that a control group would receive the 'usual' input and that the effect of the novel intervention would be compared with that usual intervention. However, while this may be feasible within one educational institution, it would be problematic across several institutions. What constitutes 'usual' input is likely to differ across institutions. It is also important to consider the complexity of clinical learning in nursing education, that is affected by a physical environment, psychosocial and relational factors, organisational culture and teaching and learning components in the clinical settings (Flott and Linden, 2016). Even if it would be possible ethically and methodologically to set a proper control group of students, it would be extremely challenging to control the contextual variables where the intervention will take place (e.g. the organisational culture, the mentoring competencies, ward staff skill-mix).

Randomisation

It is hard to dissociate issues of control from those associated with randomisation. Randomisation is used to allocate participants to either control or intervention groups in between-subjects designs. The purpose of randomisation is to minimise bias in allocating participants to study groups and, depending on the specific experimental design being used, there are different approaches to randomisation. Whatever method of randomisation is achieved it is used to prevent researchers deciding who is allocated to which study group and to prevent self-selection to study groups by participants. However, simple randomisation whereby any given participant has an equal and independent chance of being allocated to a study group can rarely be achieved in educational research as large classes may be divided across the timetable and there may be established study groups within sections of classes.

Ethical aspects of controlled experiments

The ethical aspects of controlled experiments were referred to above. The point was made that, similarly to clinical studies, it would be unethical to withhold education from one group to test the effectiveness of a novel intervention, it would be equally unethical to persist with an educational model in one group if it was known to be poor or even detrimental. The issue does no arise to the same extent, although ethical research principles still apply, to within-subjects studies because the intention at some point in the study is to provide the novel intervention. Therefore, simply because a study is educational and not clinical is not an excuse to relax the ethical aspects of the design and the intention should always be to promote the principles espoused by Beauchamp and Childress (2019) of: autonomy; non-maleficence; beneficence; and justice. Clearly, if educational experiments use patient outcomes as a measure of effectiveness then the ethical aspects related to the patients need to be considered in addition to those related to the educational aspects and these studies should be registered on a recognised clinical trials registry (Dickersin & Rennie, 2003).

Cluster randomised trials

Where randomisation is not feasible within classes or even within one educational institution then, if cooperation can be achieved across several institutions it may be possible to establish cluster randomised trials (Hemming et al., 2017). In cluster randomised trials, the individual participants are not randomised. Instead, the units to which participants belong are randomised. This approach is common in clinical research and has been applied to educational research and there is evidence, although not widespread, that this has been applied to nursing education. Most commonly it appears to have been applied to nursing education in practice, in other words across clinical centres, with Registered Nurses rather than with nursing schools; for example, by Abraham et al., (2019).

Quasi-experimental designs

As discussed above, the main features of between-subjects research (blinding; control; and randomisation) may not be as easily achieved in educational research as they are in clinical research. In what follows, we wish to illustrate how between-subjects research can be applied to educational research in nursing with reference to a range of designs and frameworks.

If investigators compromise on any of the cardinal features of an experiment, as outlined above, either blinding, control or randomisation then they are conducting a quasi-experiment (Harris et al., 2006).

As explained above, in educational research achieving any of these features is hard and achieving all three is even harder, therefore, a great deal of educational research in nursing using between-subjects designs proceeds using quasi-experimental designs.

Where a true experiment is not possible then quasi-experiments are both feasible and necessary. However, investigators must describe quasi-experiments as such and explain which feature or features make the study a quasi-experiment and demonstrate an understanding of what the consequences are. For example, if randomisation is not possible then the possibility of bias in the allocation of participants to study groups arises. If blinding is not possible then there is the likelihood of bias in measuring the outcomes of the study. If control is not possible then the study is a within-subjects design. We raise these points because we have witnessed in our work as editors of nursing education journals that some authors either do not understand these points or, if they do, fail to explain them clearly.

Alternative trial designs

Where it proves impossible to conduct traditional between-subjects experiments with full control, blinding and randomisation, nurse education researchers should consider some relatively recently developed alternatives which are being widely applied in clinical research. The fact that these are alternatives does not, necessarily detract from their rigour. Amongst these methods are non-inferiority trials and pragmatic trials and these will be considered, in turn, below.

Non-inferiority trials

'Noninferiority trials test whether a new experimental treatment is not unacceptably less efficacious than an active control treatment already in use' (Hahn, 2012; p.403) and are becoming increasingly common in clinical research. This type of design could be ideal for nursing education and, to some extent, overcomes one of the issues raised above of deciding on an appropriate control group for a study. Novel interventions can be delivered and compared with existing delivery on the basis of non-inferiority provided that the non-inferiority criteria are set appropriately and in advance of conducting the study. While these studies are well represented in clinical nursing research, they are less adopted in educational research. An example in clinical research is Boman et al., (2021) on the implementation of advanced practice nurses on orthopaedic care but this design has been used to test educational interventions with nurses performing preoperative assessment (Kinley et al., 2002).

Pragmatic trials

Patsopoulus (2011) describes pragmatic trials as follows, they are: ‘designed to evaluate the effectiveness of interventions in real-life routine practice conditions, whereas explanatory trials aim to test whether an intervention works under optimal situations.’ (p.217). Therefore they: ‘can be generalized and applied in routine practice settings.’. This approach seems eminently suitable for research into nursing education and has, indeed, been used to test the effectiveness of video education for staff working in nursing homes as reported by Mor et al. (2017). Pragmatic trials are designed to evaluate the effectiveness of interventions; they have higher external validity than explanatory trials. The latter use relatively small samples under highly artificial conditions, whereas pragmatic trials use large samples and, instead of controlling for all possible variables, test interventions against a background of high variability, which is more realistic.

A framework, leading to a tool to assess the extent to which a trial is pragmatic—the PRECIS tool (Thorpe et al., 2009)–has been developed and the features of the original framework, adapted from Patsopoulus (2011) will be used below to show how a trial may be pragmatic and how this may be applied to educational research in nursing. A wholly pragmatic trial will have no inclusion or exclusion criteria and there will be a very flexible approach to applying the intervention which is applied across all the settings that are available. For the purposes of comparison, the most suitable available alternatives to the intervention will be used and, likewise, this will be applied across all the settings that are available. The outcome that is measured will be the one that is most meaningful to the purpose of the study and will not require trained specialist data collectors. Those applying the intervention will not be trained or followed up to adhere to the trial protocol. Participants will not be followed up formally, but all potential participants will be included, and intention-to-treat analysis will be applied.

It can immediately be seen that the above framework is suitable to educational research in nursing and could be applied across a range of settings relatively easily. Of course, if it is possible to be less pragmatic in any features of the study then this should be attempted and the extent to which this has been achieved can be measured using a modified version of the PRECIS tool called the PRECIS-2 tool (https://www.precis-2.org/; accessed 2 February 2022).

Complex interventions

Complex interventions, developed for and increasingly popular in medicine, are being used in nursing research generally (Richards & Borglin, 2011) and may be suitable for investigating educational interventions. According to Petticrew (2011; p.397): “There are many other definitions of complex interventions. These tend frequently to emphasize that they have multiple interacting components, and non-linear causal pathways. Complex interventions are often contrasted in the health literature with ‘simple’ interventions, in particular medical interventions, which are generally seen as having simple linear pathways linking the intervention and its outcome.”

A complex intervention encloses multi-faceted properties of the intervention itself, the behavioural variability of the participants and the setting, and the expertise required to deliver the intervention (Skivington et al., 2021). According to the framework for developing and evaluating complex interventions, it is essential to consider the context (1), develop, refine, and test the theory behind the intervention (2), engage the stakeholders (3), identify the key uncertainties (4), refine the intervention (5), and include economic considerations (6) (Skivington et al., 2021).

Nursing education research is complex by definition, especially when considering the clinical learning environments, which are affected both by the complexity of learning and of the organisational setting (Tomietto, 2018). Clinical learning involves many different variables related to the patients (Berndtson et al., 2019), the staff members, the ward manager (Tomietto et al., 2022), and the mentors’ competencies (Mikkonen et al., 2022). All these variables deeply affect the study designs, and this complexity reflects the major challenge of implementing rigorous experimental designs in nursing education research, especially when considering the clinical learning environments.

There is no evidence that these have been used to investigate nursing education but, given the multiple components involved in educational interventions, including the content, the environment, and mode of delivery a complex intervention framework could be used in nursing education research. Complex interventions are not instant solutions to deriving answers. They require investigators to recognise that their interventions have complexity, to explain which part of the complex intervention continuum that are investigating (MRC, 2021), to show how they are addressing complexity and how they are—if possible—isolating cause and effect.

Discussion

Against a background of a paucity of rigorous experimental research in nursing education, we consider above the reasons why this is the case. There are many valid reasons why research in nursing education often cannot apply the rigorous and relatively simple experimental approaches used, for example, in randomised controlled trials. However, there is increasing recognition in clinical research that the classical randomised controlled trial is not possible, the outcomes not necessarily applicable in clinical practice and that alternative approaches are required. This has, for example, been encapsulated under the umbrella of 'real world research' (Price et al., 2015).

As such, we particularly recommend that research in nursing education considers implementing cluster randomised trials, non-inferiority trials and pragmatic trials and that these are considered within a complex intervention framework. Therefore, we would like to see researchers in nursing education consider how these designs could be used to establish if educational interventions are effective and make more significant contributions to the body of knowledge in our field. This will necessitate a more profound understanding of alternative designs and the application of rigorous standards in reporting experimental work in nursing education.

We further recommend that in any type of experimental study, the theoretical framework is carefully planned and developed. Researchers need to build their experimental study on previous evidence, clearly defined concepts, intervention content and modelling process. A health economic perspective should be also promoted in evaluating the interventions in nursing education, as a part of the complex interventions framework.

In addition, while few nursing education studies will be eligible for registration on clinical trials registries—unless the outcomes of the studies are clinical—we strongly recommend that researchers in our field consider the application of open science approaches to their research and publication in parallel with these approaches in clinical research. Therefore, we specifically suggest that study protocols in nursing education are registered on open science framework websites and study databases shared there along with links to publications. We also suggest that more use is made of pre-printing websites and to draw readers' attention to the fact that MedRxiv has a specific category for nursing research.

Conclusion

Having reviewed a range of experimental designs, we explain why within-subjects studies possibly dominate in nursing education research. Within-subjects studies are ideal for preliminary testing of interventions, and they are very convenient to conduct. However, to test educational interventions more rigorously, we urge nursing educational researchers to explore more rigorous designs and to consider using more between-subjects designs including RCTs and non-inferiority trials. Once efficacy of nursing educational interventions has been established their effectiveness can subsequently be studied in real world situations using pragmatic designs.

References

Baldi, I., Dal Lago, E., De Bardi, S., Sartor, G., Soriani, N., Zanotti, R., & Gregori, D. (2014). Trends in RCT nursing research over 20 years: mind the gap. *British Journal of Nursing*, 23, doi: 10.12968/bjon.2014.23.16.895.

Beauchamp, T.L., Childress, J.E. (2019). *Principles of biomedical ethics 8th edition*, Oxford University Press, Oxford.

Bell, M. L., Kenward, M. G., Fairclough, D. L., & Horton, N. J. (2013). Differential dropout and bias in randomised controlled trials: when it matters and when it may not. *BMJ*, *346*: e8668.

Berndtsson, I. E., Karlsson, M. G., & Rejnö, Å. C. (2019). Nursing students' attitudes toward care of dying patients: A pre-and post-palliative course study. *Heliyon*, *5*(10), e02578.

Boman, E., Duvaland, E., Gaarde, K., Leary, A., Rauhala, A., & Fagerström, L. (2021). Implementation of advanced practice nursing for minor orthopedic injuries in the emergency care context: a non-inferiority study. *International Journal of Nursing Studies*, *118*, 103910.

Campbell, D. T., Stanley, J. C., & Gage, N. L. (1963). *Experimental and quasi-experimental designs for research.* Houghton, Mifflin and Company.

Carbogim, F., Barbosa, A. C. S., de Oliviera, L. B., de Sá Diaz, F. B. B., Toledo, L. V., Alves, K. R., Friedrich, D.B.C., Luiz, F.S. & de Araújo Püschel, V. A. (2018). Educational intervention to improve critical thinking for undergraduate nursing students: A randomized clinical trial. *Nurse Education in Practice*, *33*, 121-126.

Dickersin, K., Rennie D, (2003). Registering Clinical Trials. *JAMA*, 290, 516-523.

Flott, E. A., & Linden, L. (2016). The clinical learning environment in nursing education: a concept analysis. *Journal of Advanced Nursing*, *72*(3), 501-513.

Hagen, B., Awosoga, O., Kellett, P., Ofori Dei, S. (2013) Evaluation of undergraduate nursing students' attitudes towards statistics courses, before and after a course in applied statistics. *Nurse Education Today*, 33, 9, 949-955.

Hahn, S. (2012) Understanding noninferiority trials. *Korean Journal of Paediatrics*, 55, 403-407.

Harris, A.D., McGregor, J.C., Perencevich, E.N., Furuno, J.P., Zhu, J., Petersen, D.E., & Finkelstein, J (2006). The Use and Interpretation of Quasi-Experimental Studies in Medical Informatics. *Journal of the American Informatics Association*, 13, 16-23.

Hemming, K., Elridge, S., Forbes, G., Weijer, C., & Taljaard, M. (2017). How to design efficient cluster randomised trials. *BMJ*, 358, j3064.

Keogh-Brown, M.R., Bachmann, M.O., Shepstone, L., Hewitt, C., Howe, A., Ramsay, C.R., Song, F., Miles, J.N., Torgerson, D.J., Miles, S., Elbourne, D., Harvey, I., Campbell, M.J. (2007). Contamination in trials of educational interventions. *Health Technology Assessment*, 11(43), iii, ix-107.

Kinley, H., Czoski-Murray C., George, S., McCAbe, C., Primrose, J., Reilly, C., Wood, R., Nicolson, P., Healy, C., Read, S., Norman, J., Janke, E., Alhameed, H., Fernandes,. N., Thomas, E., on behalf of the Op-Check Study Group (2002) Effectiveness of appropriately trained nurses in preoperative assessment: randomised controlled equivalence/non-inferiority trial *BMJ*, 325, 1323.

Long, T. & M. Johnson (2002). Research in Nurse Education Today: do we meet our aims and scope? *Nurse Education Today,* 22: 85-93.

Medical Research Council. (1948), Streptomycin treatment of pulmonary tuberculosis: a Medical Research Council investigation. *BMJ*, 2, 769–82.

Medical Research Council (2021). *Developing and evaluating complex interventions*, MRC, London.

National Institute of Nursing Research. (2022). *Congressional justification FY 2022*, Department of Health and Human Services, National Institutes of Health

Patsopoulos, N. (2011). A pragmatic view on pragmatic trials. *Dialogues in Clinical Neuroscience*, 13, 217-224.

Mikkonen, K., Tomietto, M., Tuomikoski, A. M., Miha Kaučič, B., Riklikiene, O., Vizcaya-Moreno, F., ... & Kääriäinen, M. (2022). Mentors' competence in mentoring nursing students in clinical practice: Detecting profiles to enhance mentoring practices. *Nursing Open*, *9*, 593-603.

Mor, K., Volandes, A., Gutman, R., Gatsonis, C., Mitchell, S. (2017). PRagmatic trial Of Video Education in Nursing homes: The design and rationale for a pragmatic cluster randomized trial in the nursing home setting. *Clinical Trials*, doi: 10.1177/1740774516685298

Morton, P.G. (2017). Nursing education research: an editor's view. *Journal of Professional Nursing*, 33, 311-312.

Padilha, J. M., Machado, P. P., Ribeiro, A., Ramos, J., & Costa, P. (2019). Clinical virtual simulation in nursing education: randomized controlled trial. *Journal of medical Internet research*, *21*(3), e11529.

Petticrew, M. (2011). When are complex interventions 'complex'? When are simple interventions 'simple'? *European Journal of Public Health*, 21, 397–398.

Price, D., Brusselle, G., Roche, N., Freeman, D., & Chisholm, D. (2015) Real-world research and its importance in respiratory medicine. *Breathe*, 11, 26-38.

Richards, D. A., & Borglin, G. (2011). Complex interventions and nursing: looking through a new lens at nursing research. *International journal of nursing studies*, *48*(5), 531-533.

Schneider, B. S., Nicholas, J., & Kurrus, J. E. (2013). Comparison of methodologic quality and study/report characteristics between quantitative clinical nursing and nursing education research articles. *Nursing education perspectives*, 34, 292–297.

Skivington, K., Matthews, L., Simpson, S. A., Craig, P., Baird, J., Blazeby, J. M., Boyd, K.A., Craig, N., French, D.P., McIntosh, E., Rycroft-Malone, J., Whjite, M., & Moore, L. (2021). A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ*, *374*.

Steckler, A., McLeroy, K.R. 2008). The importance of external validity. *American Journal of Public Health*, 98, 9-10.

Tamaki, T., Inumaru, A., Yokoi, Y., Fujii, M., Tomita, M., Inoue, Y., Kido, M., Ohnno, Y., & Tsujikawa, M. (2019). The effectiveness of end-of-life care simulation in undergraduate nursing education: A randomized controlled trial. *Nurse Education Today*, *76*, 1-7.

Thorpe, K.E., Zwarenstein, M., Oxman, A.D., Treweek, S., Furberg, C.D., Altman, D.G., Tunis, S., Bergel, E., Harvey, I., Magid, D.J., Chalkidou, K. (2009). A pragmatic-explanatory continuum

indicator summary (PRECIS): a tool to help trial designers. *Journal of Clinical Epidemiology*, 62, 464-475.

Tomietto, M. (2018). A good clinical learning environment as an organizational challenge. In Saarikoski, M., Strandell-Laine, C.(Eds.) *The CLES-scale: An evaluation tool for healthcare education* Springer, Cham. pp. 57-70.

Table 1 Experimental designs

Design	Main features
Within-subjects	The same participants are studied before and after and intervention
Between-subjects	At least two groups of participants are compared one of which receives no intervention (if not an RCT it is described as a quasi-experiment)
RCT	A between-subjects design where participants are randomly assigned to groups, one of which is a control and selected according to strict inclusion and exclusion criteria
Non-inferiority	An RCT type study where the aim is to ensure that a novel intervention is no worse than an existing treatment
Pragmatic	A real world experiment where groups of participants are treated or assigned to a control group with few inclusion or exclusion criteria