

Evaluation of hospital palliative care teams: strengths and weaknesses of the before-after study design and strategies to improve it

S Simon Institute of Palliative Care, Oldenburg, Germany; King's College London, School of Medicine, Department of Palliative Care, Policy and Rehabilitation, London and **IJ Higginson** King's College London, School of Medicine, Department of Palliative Care, Policy and Rehabilitation, London

Hospital palliative care teams (HPCTs) are well established as multi-professional services to provide palliative care in an acute hospital setting and are increasing in number. However, there is still limited evaluation of them, in terms of efficacy and effectiveness. The gold standard method of evaluation is a randomised control trial, but because of methodological (e.g., randomisation), ethical and practical difficulties such trials are often not possible. HPCT is a complex intervention, and the specific situation in palliative care makes it challenging to evaluate (e.g., distress and cognitive impairment of patients). The quasi-experimental before-after study design has the advantage of enabling an experimental character without randomisation. But this has other weaknesses and is prone to bias, for example, temporal trends and selection bias. As for every study design, avoidance and minimisation of bias is important to improve validity. Therefore, strategies of selecting an appropriate control group or time series and applying valid outcomes and measurement tools help reducing bias and strengthen the methods. Special attention is needed to plan and define the design and applied method. *Palliative Medicine* (2008); 00: 1–6

Key words: evaluation studies; health service; palliative care; patient care team; research design

Introduction

There is an ongoing debate about how to improve methodological quality in palliative care research.^{1,2} In addition, it is an ethical and methodological challenge to carry out research in the sensitive field of palliative care.^{3,4} A recent summary showed an increasing activity in scientifically based aspects of palliative care; the quantity of trials and scientific papers were steadily rising.⁵ However, the methodological quality for improving evidence is still weak compared with other specialities like oncology or cardiology. There are only a few randomised control trials (RCTs) evaluating of palliative care services. Some of these trials failed to complete because of methodological problems.^{6,7}

Hospital palliative care teams (HPCTs) are widely used to improve quality of care for patients with life-threatening illnesses within a hospital through all departments. Since the first HPCT was founded in 1976, the number of teams in hospitals has increased, both in the

United Kingdom and worldwide (in 1999, there were 340 HPCTs in United Kingdom).⁸ However, there is still a lack of scientific evidence that a HPCT will improve the quality of care, and methodological problems arise in the evaluation of complex interventions like a HPCT.⁸

The aim of this review is to describe the difficulties in the evaluation of a HPCT as a complex intervention and to discuss pros and weaknesses of the before-after study design.

Evaluation of an intervention

In the evaluation of an intervention, two different objectives can be pursued: determining the intervention's efficacy or effectiveness. Efficacy is the measurement of the intervention's potential benefit under ideal circumstances, asking whether an intervention works in principle. Effectiveness trials look at the potential benefits under real conditions in everyday practice, assessing whether the intervention has an effect for a specific population. Distinguishing between determining efficacy or effectiveness in a trial is important because the study design is defined by the trial's objective. Explanatory trials (to determine efficacy) use experimental designs, like a RCT, while

Correspondence to: Dr Steffen Simon, Visiting Research Associate, King's College London, School of Medicine, Department of Palliative Care, Policy and Rehabilitation, Weston Education Centre, Denmark Hill, London SE5 9RJ, UK.
Email: steffen@steffensimon.de

pragmatic trials (to determine effectiveness) can also be randomised trials but can also include observational studies.⁹ For example, in designing a new drug, it is necessary to prove the efficacy in RCTs in highly selected populations first, and afterwards its effectiveness can be studied in multicentre trials and in daily practice with a cohort study.

There are three different types of epidemiological study designs which can be used for evaluation of interventions (Table 1).¹⁰

Observational studies (cross sectional or cohort) explore the association between a potential causative factor (the independent or exposure variable) and the outcome of interest (dependent or outcome variable).¹¹ For example, a symptom improvement in cancer patients after implementation of a HPCT could be observed in an observational study.¹² However, cross-sectional studies are less eligible to explore the association as data are collected only at one time. Cohort studies are more useful in this situation.

Experimental studies are the only design to assess causal relationships. There are different types of RCT: classical ones like parallel groups, crossover and cluster randomisation and, more recently, developed methodologies like a delayed intervention RCTs or a patient preferences design.

The quasi-experimental design is interventional. Because there is no randomisation, selection bias becomes an issue. In general, its limitations are more related to observational studies.¹³ The before-after design is the most widely used quasi-experimental design, but it is very weak without a control group and time series data.

Evaluation of a HPCT

A HPCT is a multi-professional service that provides specialist palliative care through advisory, educational and direct care to patients and their families within the acute hospital setting. The team usually includes a specialist nurse and a consultant in palliative medicine with support from social workers, chaplains and pharmacists.¹⁴ Even

though the usefulness of a HPCT seems to be obvious, there is still a lack of evidence regarding efficacy and effectiveness. A systematic review in 2002 found only nine (plus four with broader intervention) evaluation studies of a HPCT, which show a small benefit and a wide range of methodological problems.⁸ There was a 'paucity of good evaluation, of any comparative design, in this field', and further, well-designed studies with good validity are needed.

Most of the evaluation studies of HPCTs are effectiveness trials with observational designs,^{8,12,14-18} as previously mentioned, only two are RCTs.^{19,20}

Difficulties in evaluation of HPCT and in palliative care

However, there are multiple difficulties and pitfalls for an evidence-based approach in evaluating a HPCT.^{2,6,21,22} First, the complex nature of a HPCT as a health service is one of the reasons for difficulty in designing and evaluating such an intervention (Table 2).

The complex nature of a HPCT derived from many reasons, such as different professionals and different investigations. Regarding a team with nurses, physician and social workers, a study has to take into account the different backgrounds, education, responsibilities and functions of the team members.²³ Also, the study design must consider the different kinds of intervention delivered by a HPCT (e.g., drugs and communication skills), how they influence the outcome and how to measure them. Because of heterogeneity regarding settings, experiences, training, etc. and lack of standardisation, it is very difficult to compare different HPCTs; hence the need for careful definition.⁸

Second, difficulties in evaluation arise because of the sensitive field of palliative care, leading to many problems for conducting a study (Table 3).

Patients with palliative care needs are a very vulnerable patient group with some delicate characteristics and difficulties for research. Heterogeneity in diagnoses and stages of diseases make standardisation and definition difficult,

Table 1 Types of epidemiological study designs

Types of study	Artificial manipulation of the study factor	Randomisation	Study designs
Observational	No	No	Cross sectional or cohort
Quasi-experimental	Yes	No	Before-after
Experimental	Yes	Yes	RCT

RCT, randomised control trial.

Amended Table from: Constantini M and Higginson IJ. Experimental and quasi-experimental designs. In: Addington-Hall JM, *et al.* Research Methods in Palliative Care. 1ed. Oxford: Oxford University Press; 2007. Page 86, Table 6.1 (by permission of Oxford University Press).

Table 2 Complex nature of health services (e.g. HPCT)

Health service	HPCT
Different professionals	Physicians, nurses, social workers, volunteers, etc.
Different investigations	Medication, education, support, communication, etc.
Different settings and organisations	Public or private hospitals, etc.
Different components in outcomes and outcome measurements	Quantitative and qualitative data; lack of valid measurement tools in palliative care
Difficult to standardise the intervention for comparison and replication	Differences in settings, experiences, training, personal characteristics, commitment, etc.

HPCT, hospital palliative care team.

along with finding appropriate control groups and avoiding contamination.¹⁹ The high attrition rate and the very limited life expectancy combined with a stressful situation and time at the end of life lead to difficulties in recruitment and often to a small sample size.⁶ Patients with cognitive impairment and the large variation of disease trajectories with rapid deterioration, changes and fluctuating course often result in substantial loss of data.² The holistic approach of palliative care and its services cause some problems in defining clear outcomes and finding valid measurements.² Table 3 summarises the problems in the field of palliative care, which are discussed in depth elsewhere, for example, Grande and Todd give a comprehensive overview.²²

Randomisation

Randomisation is the allocation of individuals to groups (e.g., intervention and control group) by chance to ensure similarity between the groups.²⁴ If the sample size is large enough, randomisation is the most effective way to distribute all confounders equally between the groups to relate the differences in the results to the intervention alone. From an ethical point of view, neither interventions (neither the new intervention nor the standard treatment) should be proven superior; there must be scientific uncertainty (so called 'equipoise').

A recent non-systematic review found only a few randomised studies in palliative care (22/539 total studies).⁵

Two older systematic reviews explore 11 RCTs in palliative care with various interventions (service, consultations, chemotherapy, etc.) of which two failed to complete.^{2,25} The main reason for the opposition of patients, carers and professionals (referral source) to randomisation is the belief that allocation by chance is unethical for patients at the end of their life. Many people believe (non-scientifically) that, for example, being cared by a new HPCT will be better than without such care. Therefore, 'equipoise' does not exist and randomisation is refused either by the participants or by those referring them into the study.⁶ The priorities of professionals for further care of the patient provide another pressure against allocation by chance. Providing a delayed intervention (the control group with a waiting list achieves the new intervention as well but later) has been proposed as an alternative to maintain randomisation and encourage participation.²⁶ One of the RCTs in evaluation of HPCT used a cluster randomisation by randomising practices with their registered patients and not the patients directly.¹⁹ Using this approach, they recruited a large number of patients in their trial. The other RCT in evaluation of HPCT used an unequal ratio (2:1) for randomisation to improve recruitment into the intervention arm.²⁰ However, only 38% of potential patients participated and the randomised groups were not representative when compared with the non-randomised group; non-participants were less well and had a shorter survival time. Both studies failed to demonstrate a distinct improvement following the intervention. In conclusion, the evaluation of a palliative care service with a RCT is

Table 3 Difficulties in the field of palliative care

Problem	Examples in palliative care
Failure in recruitment	Opposition to randomisation by patients and professionals
Difficulties in selection	Heterogeneity of palliative patients, attrition, instable course of disease at the end of life
Ethical problems	Randomisation of dying patients
Difficulties in measurement	Distress and cognitive impairment of palliative patients in completing questionnaires or interviews
Difficulties in design	Appropriate control group, complex outcomes like quality of life or satisfaction, lack of validity of measurement tools
Contamination	Between intervention and control group in the process of care, influences of personal characteristics of the team members or the researcher, problems with blinding

complicated by several dilemmas and much more challenging than the investigation of a drug.

Two alternatives might serve as a solution to overcome this problem. On the one hand, the Medical Research Council framework for the evaluation of complex interventions was developed to help researchers to design a new complex intervention like a health service, ensuring high internal and external validity by usage of RCTs and to attend them throughout the research process.²¹ One of us is currently using this framework in the field of palliative care to assess a new service for patients with multiple sclerosis and palliative care needs²⁶ and for patients with intractable dyspnoea.²⁷ The framework is discussed elsewhere.²⁸

In contrast, if a RCT is not feasible or possible to conduct, a well-designed, quasi-experimental study like the before-after design might be an alternative in evaluating palliative care services.

Before-after study design

The design is characterised by two measurement points or periods of time: one before and one after the intervention. Therefore, it is called a pretest-posttest design. In the classical design, two groups are compared on an institutional level, for example, the quality of care of dying patients in a general hospital is compared before and after the implementation of an inpatient palliative care program during a period of 6 months.²³ A control group is essential to strengthen the validity.²⁹ The before-after study design is mainly used for interventions like educational programs,³⁰ implementation of guidelines,³¹ assessment tools,³² health services²³ or new orders and laws.³³

However, there are pros and weaknesses of the before-after study design.

The greatest advantage is the experimental character without randomisation. Therefore, it is possible to include all patients at one institution in the experimental group so that they can all achieve the new service or intervention. For example, setting up a new care program and including all patients who need the new service and evaluating the service by comparing a measurement before and after the implementation.²³ Furthermore, the before-after study design is more feasible and much less expensive and time consuming compare with a RCT.

However, limitations of the design are responsible for the low statistical significance of the association between the intervention and its outcomes, leading to only very cautious interpretations and conclusions. The internal validity of the before-after design is weakened by some specific and general methodological problems.

First, *secular* (or *temporal*) *trends* mean a change over a period of time as a general development independently

from the intervention, which could itself explain the change in the measured outcomes.²⁴ Secular trends are estimated as ‘potentially the greatest threat to the validity of before-after studies’.³³ As an example, the evaluation of an intervention to reduce terminal hospitalisations would have to take into account the general trend of hospitalisations near to death.³⁴ The most frequently used before-after design without control group is very vulnerable to this confounder in not allowing for control over this external effect.

Second, *regression to the mean* is the tendency of individuals at the extreme or outliers to have values nearer to the mean of the general population on repeated measurements.²⁴ For example, without any intervention, a very low level of palliative care knowledge tends to improve solely as a consequence of general improvement in palliative care knowledge.¹⁷

Third, *bias* is the deviation of results from the truth as a systematic error.²⁴ It is an incorrect estimate of the outcome of interest. A lot of different types of bias exist, but they can be classified mainly in two groups: selection bias and measurement bias (synonym information or observational bias). Selection bias is the systematic difference in characteristics between those who take part in a study and those who do not, for example, the difference in stages of the disease after a non-representative recruitment of only 38% of potential patients in evaluation of a HPCT.²⁰ Measurement bias arises from inaccurate measurements or classification of subjects or outcomes, for example, to define and measure the quality of end-of-life care by determining the documented order of opioid medication.²³

Fourth, *confounding* is a situation in which a measure of the effect of an exposure on risk is distorted because of the association of exposure with another factor that influences the outcome under study.²⁴ Confounders are extraneous variables that, if uncontrolled for, affect the dependent variable, for example, the setting up of a different palliative care service within the study area by an external charity while evaluating a coordinating care service for terminally ill cancer patients, which substantially influences the control group.¹⁹

Methodological implications and strategies

How to deal with the different weaknesses of the before-after study design?

Diverse strategies can be applied to strengthen the validity, useful for the before-after study design, but feasible for all designs in evaluation of HPCTs (see Table 4). Randomisation is the most effective strategy to overcome all the potential weaknesses in the before-after study design.

In the analysis of a before-after study, it is also necessary to detect all potential bias³⁶ and to assess the extent

Table 4 Checklist of methodological strategies

Potential weakness	Is there a risk that the change over time could be explained by...	Strategies to reduce weakness – what to do?
Secular trend	...the general development independent from the intervention?	Control group or time series ²⁹ Minimise the duration of monitoring
Regression of the mean	...the tendency of individuals at the extreme to have values nearer the mean of the population on repeated measurements?	Control group or time series
Selection bias	...differences in characteristics between those who take part in a study and those who do not?	Comparable groups 'before' – 'after'
Measurement bias	...systematic error measurements or classification of subjects or outcomes?	Equal periods of monitoring, in length (e.g., 6 months) and chronology (January–June) ³⁴ Standardised assessment and valid measurement tools (e.g., STAS* or audits) ^{16,35} Prospective approach Blinding measurements
Confounding	...an extraneous variable that affects the dependent variable?	Matching (by age, gender, education, etc.) Restriction (excludes known confounders)

* STAS, support team assessment schedule.

and direction of bias if possible. In the discussion part, the bias and confounders with their direction of influence should be considered with reference to the hypothesis. Stratification of the sample (e.g., in age groups) allows reduction in confounding. There are some other methods to avoid systematic errors like the 'intention to treat' analysis or the 'multivariate analysis'.

Conclusion

In palliative care research, there is a need for more rigorous studies to improve evidence. In agreement with Kaasa, *et al.*⁵ we would like to encourage clinicians and researcher to discuss methodological issues frequently and collaborate on debating pitfalls and problems in research. We recommend using a RCT to evaluate an intervention whenever it is appropriate and feasible because it is still the most effective way to control and reduce bias. As discussed, in particular for the evaluation of a complex intervention in the field of palliative care, like a HPCT, the before-after study design could be an appropriate alternative. However, this design has limitations that need extra precautions in conducting and interpretation of the study. Therefore, some specific strategies (i.e., having a control group, time series, valid outcome measures, etc.) to control and minimise secular trends, bias and confounders are required, which will help to strengthen the validity in further research of this approach.

Acknowledgment

The authors thank Christina Ramsenthaler, Claudia Bausewein and Fliss Murtagh, who read former drafts of this article.

References

- Higginson, IJ. Evidence based palliative care: there is some evidence-and there needs to be more. *BMJ* 1999; **319**: 462–463.
- Rinck, GC, van den Bos, GA, Kleijnen, J, de Haes, HJ, Schade, E, Veenhof, CH. Methodologic issues in effectiveness research on palliative cancer care: a systematic review. *J Clin Oncol* 1997; **15**: 1697–1707.
- Addington-Hall, JM. Research sensitivities to palliative care patients. *Eur J Cancer Care* 2002; **11**: 220–224.
- Kendall, M, Harris, F, Boyd, K, Sheikh, A, Murray, SA, Brown, D, *et al.* Key challenges and ways forward in researching the "good death": qualitative in-depth interview and focus group study. *BMJ* 2007; **334**: 521.
- Kaasa, S, Hjerstad, MJ, Loge, JH. Methodological and structural challenges in palliative care research: how have we fared in the last decades. *Palliat Med* 2006; **20**: 727–734.
- McWhinney, IR, Bass, MJ, Donner, A. Evaluation of a palliative care service: problems and pitfalls. *BMJ* 1994; **309**: 1340–1342.
- Higginson, IJ, Finlay, IG, Goodwin, DM, Hood, K, Edwards, AG, Cook, A, *et al.* Is there evidence that palliative care teams alter end-of-life experiences of patients and their caregivers? *Journal of Pain & Symptom Management* 2003; **25**: 150–168.
- Higginson, IJ, Finlay, I, Goodwin, DM, Cook, AM, Hood, K, Edwards, AG, *et al.* Do hospital-based palliative teams improve care for patients or families at the end of life. *J Pain Symptom Manage* 2002; **23**: 96–106.
- Bausewein, C, Higginson, IJ. Appropriate methods to assess the effectiveness and efficacy of treatments or interventions to control cancer pain. *J Palliat Med* 2004; **7**: 423–430.
- Constantini, M, Higginson, IJ. Experimental and quasi-experimental designs. In: Addington-Hall, JM, Bruera, E, Higginson, IJ, Payne, S, (eds), *Research methods in*

- palliative care. 1st ed. Oxford: Oxford University Press; 2007. p. 85–113.
- 11 Black, N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996; **312**: 1215–1218.
 - 12 Morita, T, Fujimoto, K, Tei, Y. Palliative care team: the first year audit in Japan. *J Pain Symptom Manage* 2005; **29**: 458–465.
 - 13 Schulz, KF, Grimes, DA. The lancet - handbook of essential concepts in clinical research. 1st ed. London: Elsevier; 2006.
 - 14 Hunt, J, Keeley, VL, Cobb, M, Ahmedzai, SH. A new quality assurance package for hospital palliative care teams: the Trent Hospice Audit Group model. *Br J Cancer* 2004; **91**: 248–253.
 - 15 Fromme, EK, Bascom, PB, Smith, MD, Tolle, SW, Hanson, L, Hickam, DH, *et al.* Survival, mortality, and location of death for patients seen by a hospital-based palliative care team. *J Palliat Med* 2006; **9**: 903–911.
 - 16 Jack, B, Hillier, V, Williams, A, Oldham, J. Hospital based palliative care teams improve the insight of cancer patients into their disease. *Palliat Med* 2004; **18**: 46–52.
 - 17 Morita, T, Fujimoto, K, Imura, C, Nanba, M, Fukumoto, N, Itoh, T. Self-reported practice, confidence, and knowledge about palliative care of nurses in a Japanese Regional Cancer Center: longitudinal study after 1-year activity of palliative care team. *Am J Hosp Palliat Care* 2006; **23**: 385–391.
 - 18 OMahony, S, Blank, AE, Zallman, L, Selwyn, PA. The benefits of a hospital-based inpatient palliative care consultation service: preliminary outcome data. *J Palliat Med* 2005; **8**: 1033–1039.
 - 19 Addington-Hall, JM, MacDonald, LD, Anderson, HR, Chamberlain, J, Freeling, P, Bland, JM, *et al.* Randomised controlled trial of effects of coordinating care for terminally ill cancer patients. *BMJ* 1992; **305**: 1317–1322.
 - 20 Hanks, GW, Robbins, M, Sharp, D, Forbes, K, Done, K, Peters, TJ, *et al.* The imPACT study: a randomised controlled trial to evaluate a hospital palliative care team. *Br J Cancer* 2002; **87**: 733–739.
 - 21 Campbell, M, Fitzpatrick, R, Haines, A, Kinmonth, AL, Sandercock, P, Spiegelhalter, D, *et al.* Framework for design and evaluation of complex interventions to improve health. *BMJ* 2000; **321**: 694–696.
 - 22 Grande, GE, Todd, CJ. Why are trials in palliative care so difficult. *Palliat Med* 2000; **14**: 69–74.
 - 23 Bailey, FA, Burgio, KL, Woodby, LL, Williams, BR, Redden, DT, Kovac, SH, *et al.* Improving processes of hospital care during the last hours of life. *Arch Intern Med* 2005; **165**: 1722–1727.
 - 24 Last, JM. A dictionary of epidemiology. 4th ed. New York: Oxford University Press; 2001.
 - 25 Smeenk, FW, van Haastregt, JC, de Witte, LP, Crebolder, HF. Effectiveness of home care programmes for patients with incurable cancer on their quality of life and time spent in hospital: systematic review. *BMJ* 1998; **316**: 1939–1944.
 - 26 Higginson, IJ, Vivat, B, Silber, E, Saleem, T, Burman, R, Hart, S, *et al.* Study protocol: delayed intervention randomised controlled trial within the Medical Research Council (MRC) Framework to assess the effectiveness of a new palliative care service. *BMC Palliat Care* 2006; **5**: 7.
 - 27 Booth, S, Farquhar, M, Gysels, M, Bausewein, C, Higginson, IJ. The impact of breathlessness intervention service (BIS) on the lives of patients with intractable dyspnea: a qualitative phase 1 study. *Palliat Support Care* 2006; **4**: 287–293.
 - 28 MRC, Health Services and Public Health Research Board. A framework for development and evaluation of RCTs for complex interventions to improve health. London: Medical Research Council; 2000.
 - 29 Ray-Coquard, I, Philip, T, de Laroche, G, Froger, X, Suchaud, JP, Voloch, A, *et al.* A controlled “before-after” study: impact of a clinical guidelines programme and regional cancer network organization on medical practice. *Br J Cancer* 2002; **86**: 313–321.
 - 30 Back, AL, Arnold, RM, Baile, WF, Fryer-Edwards, KA, Alexander, SC, Barley, GE, *et al.* Efficacy of communication skills training for giving bad news and discussing transitions to palliative care. *Arch Intern Med* 2007; **167**: 453–460.
 - 31 Treece, PD, Engelberg, RA, Crowley, L, Chan, JD, Rubinfeld, GD, Steinberg, KP, *et al.* Evaluation of a standardized order form for the withdrawal of life support in the intensive care unit. *Crit Care Med* 2004; **32**: 1141–1148.
 - 32 Crooks, DL, Whelan, TJ, Reyno, L, Willan, A, Tozer, R, Mings, D, *et al.* The Initial Health Assessment: an intervention to identify the supportive care needs of cancer patients. *Support Care Cancer* 2004; **12**: 19–24.
 - 33 Fergusson, D, Hebert, P, Shapiro, S. The before/after study design in transfusion medicine: methodologic considerations. *Transfus Med Rev* 2002; **16**: 296–303.
 - 34 Levy, C, Morris, M, Kramer, A. Improving end-of-life outcomes in nursing homes by targeting residents at high risk of mortality for palliative care: program description and evaluation. *J Palliat Med* 2008; **11**: 217–225.
 - 35 Edmonds, PM, Stuttaford, JM, Penny, J, Lynch, AM, Chamberlain, J. Do hospital palliative care teams improve symptom control? Use of a modified STAS as an evaluation tool. *Palliat Med* 1998; **12**: 345–351.
 - 36 Delgado-Rodriguez, M, Llorca, J. Bias. *J Epidemiol Community Health* 2004; **58**: 635–641.