Issues for designing and evaluating
a 'heroin trial'.
Three discussion papers

**Feasibility Research into the Controlled Availability of
Opioids Stage 2**

Working Paper Number 8

Report on a workshop on trial evaluation
(G. Bammer and D.N. McDonald)

An evaluation of possible designs for a heroin trial
(R.G. Jarrett and P.J. Solomon)

Service provision considerations for the evaluation of a heroin trial. A discussion paper
(D.N. McDonald, G. Bammer, D.G. Legge and B.M. Sibthorpe)

Acknowledgements

Contents

Executive Summary

Three discussion papers about the design and evaluation of a trial to provide heroin to dependent users in a controlled manner are presented. They raise a number of difficult issues that need to be taken into consideration before the trial design can be finalised.

The first is a report on a workshop on trial evaluation held in June 1993. The bulk of the discussion involved the advantages and disadvantages of three evaluation designs: the randomised controlled trial design, a modified randomised controlled trial design, and the cross–over design. The types of comparisons that could be incorporated in these designs were also discussed. It was agreed that oral methadone would be the 'control' group. This could be compared with injectable heroin, or injectable heroin plus oral methadone, or with a group which had a range of choices between heroin and/or methadone and injectable, oral and smokable routes of administration. The impact of choice of study population on the generalisability of the results and on the types of outcome measures that could be used was also discussed. While there was little support for non–experimental designs, discussion of such designs clarified the limitations of the experimental approaches.

The second paper involves the statistical consideration of alternative trial designs in the case where participants can elect not to be allocated to certain treatments. It was shown that this necessarily leads to a loss of power, but that substantial improvements can be made over the 'intention to treat' analysis. Given the scenario of three possible treatments– heroin only, heroin plus methadone, and methadone only– and the likelihood that some participants will not accept allocation to treatment with methadone only, two possibilities for allocation were explored. The first was to ask, before allocation to a treatment, whether or not participants would continue to participate if they were allocated to the methadone only treatment. Those who would continue to participate would then be randomly allocated to one of the three possible groups; those who would not would be randomly allocated to either heroin only or heroin plus methadone. The second possibility would be to randomly allocate participants to one of the three arms initially, then to ask those allocated to the methadone only arm whether or not they would continue to participate. If not, they would be randomly re–allocated to either heroin only or heroin plus methadone. Statistically, the first design has clear advantages over the second. This was the design considered as the modified randomised controlled trial design in the workshop reported in the first paper.

The third discussion paper deals with potential conflicts between the provision of an effective service to participants in a 'heroin trial' and the constraints that may be imposed by the need for rigorous evaluation. In particular, the incompatibilities between a health development approach to service provision and a randomised controlled trial are explored. Two issues that are especially important, and that are a good illustration of the difficulties, are client choice and individualised treatment. A set of principles of practice is identified.

The design and evaluation of a trial of controlled provision of heroin to dependent users is an important part of the feasibility considerations. As the three discussion papers illustrate, there are many difficulties that will need to be resolved for a trial to go ahead. The final considerations about trial design and evaluation will also need to include information gathered in other components of the feasibility study

about numbers of dependent heroin users in Canberra, possible outcome measures and the preferences of potential trial participants.

Foreword

In our considerations about a trial to provide heroin in a controlled manner to dependent users, the necessity for rigorous evaluation has always been at the forefront (Douglas 1991; Bammer 1993). In our ongoing discussions about trial design and evaluation, we have also developed a number of other criteria that a trial should meet. Thus, a trial should be

- rigorous and able to give clear answers to definitive questions;
- acceptable to potential participants;
- congruent with a health development approach to service provision;
- ethical; and
- generalisable.

Designing a trial to meet these criteria is one of the most difficult aspects of the feasibility considerations. In this working paper, three discussion papers are presented, which in different ways engage some of the complex issues involved in trial design and evaluation.

In Stage 1 of the feasibility considerations, a range of design and evaluation strategies were considered (Bammer et al. 1991) and it was concluded that a randomised controlled trial was likely to be the best design for an experiment with the controlled provision of heroin to dependent users in Canberra. The 'experimental group' would have a choice of treatments between heroin and/or methadone and injectable, oral and/or smokable routes of administration. They would be compared to a group that received oral methadone only. The questions a trial would seek to address would be:

> "Can a treatment program which offers heroin (as well as methadone) and injectable and smokable routes of administration (as well as oral) increase the likelihood that participants will be able to:
>
> a.   lead a more structured lifestyle in terms of employment, relationships and day–to–day activity?
>
> b.   reduce their criminal activity?
>
> c.   reduce behaviours that place them at risk of contracting HIV and hepatitis B and C?
>
> d.   increase behaviours important in the maintenance of health and well–being?" (Bammer et al., 1991).

Potential problems with a high rate of drop–outs amongst participants allocated to the group receiving only oral methadone were recognised. This had occurred in the only previous randomised controlled trial comparing heroin and methadone maintenance treatments. In a study conducted in the late 1970s, the then new treatment option, oral methadone, was compared with injectable heroin, which was at that time the standard maintenance treatment in the United Kingdom (Hartnoll et al., 1980). Twelve percent of participants allocated to receive oral methadone dropped–out immediately; by three months less than 50 percent were still in the trial and by 12 months only 29 percent of the original participants remained in the oral methadone group. At 12 months, 74 percent of those allocated to receive injectable heroin were still in the trial. This potential problem with drop–outs piqued the interest of NCEPH statistician, Dr Patty Solomon, who began work on possible modifications to a randomised controlled trial that would allow participants some choice. When she moved to the University of Adelaide, Professor Richard Jarrett joined her in that preliminary research, the results of which are presented here as the second discussion paper.

At the same time, a small group of NCEPH and AIC researchers (Gabriele Bammer, David Legge, David McDonald and Beverly Sibthorpe) met on a monthly basis to discuss the limitations of a randomised controlled trial on the interactions between clinical trial staff and trial participants. The group was particularly influenced by the work of one of its members (David Legge) on health development, and potential incompatibilities between a health development approach to service provision in a trial and a randomised controlled trial were explored. The group read and discussed a number of articles and books, particularly about cultural attitudes to pleasure and the significance of different explanatory models of drug use in shaping models of service delivery; considered their own experiences and the experiences of guests to the group (service providers and illicit drug users) with drug treatment and prevention agencies; and developed a set of principles that might underlie the service provision in a 'heroin trial', including testing these principles in a range of scenarios. A distillation from these discussions is presented as the third paper.

Both the second and third discussion papers were provided as background reading for the June 1993 workshop. At that workshop, there was very little discussion about the issues raised in the third paper.

The design and evaluation of a trial of controlled provision of heroin to dependent users is an important part of the feasibility considerations. As the three discussion papers illustrate, there are many difficulties that will need to be resolved for a trial to go ahead. The final considerations about trial design and evaluation will also need to include information gathered in other components of the feasibility study about numbers of dependent heroin users in Canberra, possible outcome measures and the preferences of potential trial participants. We expect to be in a position to make a final decision about trial design and evaluation in May 1994 and would welcome responses to this working paper for consideration.

Gabriele Bammer

References

Bammer, G. (1993), 'Should the controlled provision of heroin be a treatment option? Australian feasibility considerations', *Addiction* 88: 467–475.

Bammer, G., Douglas, R. & Dance, P. (1991), 'Evaluation by a randomised controlled trial', in National Centre for Epidemiology and Population Health, *Feasibility Research into the Controlled Availability of Opioids Stage 1, Volume 2: Background Papers*, 279–286.

Douglas, R.M. (1991) 'Foreword', in National Centre for Epidemiology and Population Health, *Feasibility Research into the Controlled Availability of Opioids Stage 1 Volume 1. Report and Recommendations*, i–iii.

Hartnoll, R.L., Mitcheson, M.C., Battersby, A., Brown, G., Ellis, M., Fleming, P. & Hedley, N. (1980), 'Evaluation of heroin maintenance in a controlled trial', *Archives of General Psychiatry* 37:877–884.

REPORT ON A WORKSHOP ON TRIAL EVALUATION

Gabriele Bammer* and David McDonald[#]

* National Centre for Epidemiology and Population Health
# Australian Institute of Criminology

Summary

The workshop discussion focussed on five issues: the definition of the study population, the types of comparisons which could be made, possible experimental designs, outcome measures and non–experimental designs. The bulk of the discussion involved detailed consideration of three experimental designs: the randomised controlled trial design, a modified randomised controlled trial design as proposed by Jarrett and Solomon and the cross–over design. The strengths and weaknesses of each were discussed. The discussion highlighted many of the difficulties inherent in trial design and evaluation. The design and evaluation cannot be finalised until research investigating numbers of potential participants and workable outcome measures is finalised.

1 Introduction

In June 1993, a workshop about the evaluation of a trial to provide heroin to dependent users in a controlled manner was held. It aimed to build on earlier considerations about the evaluation (Bammer 1993; Bammer et al. 1991; Ostini et al. 1993) and to incorporate new research (Jarrett and Solomon, accompanying; McDonald et al., accompanying). A list of workshop participants is included at the end of this paper. The papers listed as references were provided as preliminary background reading.

The discussion was wide–ranging and fairly unstructured and, to some extent, that is reflected in this report. Very little time was spent discussing the precise research question. The gist of the question posed in Stage 1 was used as a general starting point. It was recognised that the precise question would finally be determined by an interplay between questions of interest, the evaluation design and possible outcome measures. To reiterate, the question posed in Stage 1 was:

"Can a treatment program which offers heroin (as well as methadone) and injectable and smokable routes of administration (as well as oral) increase the likelihood that participants will be able to:

a. lead a more structured lifestyle in terms of employment, relationships and day–to–day activity?

b. reduce their criminal activity?

c. reduce behaviours that place them at risk of contracting HIV and hepatitis B and C?

d. increase behaviours important in the maintenance of health and well–being?" (Bammer et al. 1991).

2 The study population – people who want treatment or people who want heroin?

It was agreed early in the workshop that the population eligible to participate in a trial needed to be specified. Dependent heroin users can be divided into those who want to participate because they are interested in treatment and those who wish to participate because they are interested in easier, safer access to heroin. Because the motivations of these two groups are different, the outcomes of their participation may also be different. In addition, choice of study population will determine how the trial is run and the generalisability of the results.

The discussion at the workshop defined the population eligible for the trial to be people who are willing to enter a treatment program for their heroin use. On this basis, it was considered that people currently on the Canberra methadone program could be the starting population to be offered a place on a 'heroin trial'. A place on a trial could also be offered to new recruits to the methadone program for a specified period of time. Thus, eligible people would be told about the trial and invited to participate. Some would refuse. Those who gave informed consent would then be randomly allocated to one of two or three groups, which would then be compared on a number of outcome measures using one of the experimental designs outlined below. The following discussion outlines groups that could be compared, potential experimental designs and potential outcome measures. There is also a brief discussion of non–experimental designs.

3 Comparisons between groups

It is possible to identify a number of comparisons between groups that could be appropriate for any of the experimental designs discussed below. They include the following:
- choice of treatments (including heroin and/or methadone with injectable, smokable and/or oral routes of administration) *versus* no choice (oral methadone only). This was the comparison proposed at the end of Stage 1 (Bammer et al. 1991);
- injectable heroin *versus* oral methadone;
- injectable heroin *versus* injectable heroin plus oral methadone *versus* oral methadone (3 groups);
- injectable heroin plus oral methadone *versus* oral methadone;
- injectable heroin plus oral methadone *versus* oral methadone *versus* injectable methadone (3 groups).

In each comparison, oral methadone is the 'control' group. It was noted that, depending on the precise nature of the research question, an arm that contains heroin only rather than a choice of treatments may have more statistical power for comparison with oral methadone because it is more homogeneous.

For the purposes of illustration, the comparison between injectable heroin, injectable heroin plus oral methadone, and oral methadone will be used in all the examples given below and the starting population will be assumed to be people currently on the methadone program.

4   Standard randomised controlled trial design

People currently on the methadone program → informed consent, then random allocation → injectable heroin / injectable heroin plus oral methadone / oral methadone; and refusal to participate.
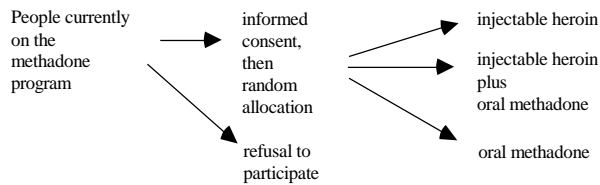
Advantages[*]

- simplicity and clarity;
- dropping–out of the trial is an outcome measure that is easy to obtain;
- compared with the cross–over trial design, people are not artificially held in the trial;
- compared with a cross–over design, there is less problem with the need to keep people in a trial for long enough to be able to unequivocally associate outcomes with treatments (see below);
- strong external validity, that is, the RCT design is considered by many people, including those who advise decision makers, to be the gold standard for evaluating a new intervention.

Disadvantages

- may not be acceptable to participants;
- if high drop–out rates are expected, it may be unethical for a trial to be established primarily to measure dropping–out of the trial as an outcome;
- if there are differences between the two groups in the numbers dropping out, there will be difficulties in measuring other outcomes that rely on self–report;
- to obtain accurate information on self–report outcomes, people who have dropped out of the trial will need to be found and followed up. This is expensive and time–consuming;
- works best with independent outcome measures, for example, crime rates measured by the police or overdose data collected in hospitals. In Australia, there are no well–established reliable data sets for either health or criminal outcome measures. In addition, obtaining such data for individuals who agree to participate in a trial may be expensive, time–consuming and difficult because of privacy considerations;
- a high drop–out rate, especially if not evenly distributed between the different groups in the trial, will adversely affect the power of the study.

Other points

There was some debate about whether or not this was the best design for generalising results to policy. Results from randomised controlled trials can be well generalised to policy if comparisons are made between communities rather than individuals, and thus where communities rather than individuals are subjected to random allocation to different 'treatments'. Lawrence Sherman pointed to his experience with studies of responses to domestic violence in this context (Sherman 1992). In the example given

---

[*] many of the advantages and disadvantages apply to more than one of the designs, but they have not always been reiterated for each design.

above, the results could be generalised to people in a treatment program with conditions similar to those in a trial, but would not be relevant to the needs of heroin users in the community generally. (This is also the case for the examples given in the other designs.)

5 Modified Randomised Controlled Trial Design

One of the designs from the accompanying paper by Jarrett and Solomon was discussed in detail. It should be noted that this design was originally conceived for a different scenario where any dependent heroin user could apply for a place on a trial– in other words, it was not restricted to people on the methadone program.
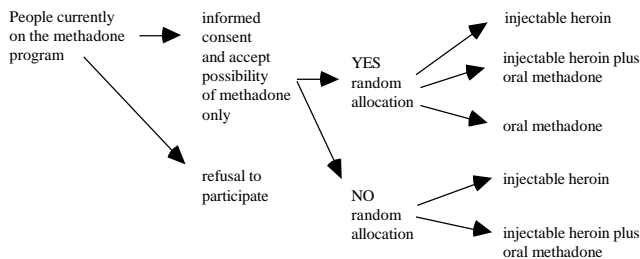
The first decision is that people would be asked if they agree to participate in an arm where they might be allocated to receive either injectable heroin only, injectable heroin plus oral methadone or oral methadone only. Essentially, this arm is the same as a standard randomised controlled trial. Participants who do not agree to participate in this arm could fall into one of two groups, those who will only participate if they receive heroin and those who will only participate if they receive methadone (this last case was not considered by Jarrett and Solomon). The first group could be assigned to heroin only or heroin plus methadone (shown above); the second group could be assigned to methadone only or methadone plus heroin (not shown).

Advantages

- offers a technical advantage over simply letting people drop–out if they do not accept the arm to which they are randomly allocated;
- fits more closely to a health development approach to service provision, because the process of allocating participants to a 'control' group is more responsive to the needs of the individual participants than a standard RCT design can be;
- likely to have better client acceptability and therefore higher recruitment and retention levels.

Disadvantages

- if a high proportion refuse to participate in the arm that is essentially the standard RCT, the trial will have low public acceptability and low generalisability (external validity);
- giving people this choice may increase the likelihood that they will not accept something they would have accepted if there was no choice;
- design rests on more assumptions than the standard RCT;
- design is more complicated than the standard RCT;

- would not allow research questions concerning retention in treatment to be addressed.

6   Cross–over design

The time period for the cross–over needs to be determined in light of the outcomes to be measured, but for the sake of this example 6 months is used. Participants would be randomly allocated to one of three groups and would stay in this group for 6 months, at which time they would swap to a second group. At the end of a further 6 months, participants would swap to a third group. The trial could finish 6 months after the swap to the third group or could have a final stage where participants were allowed a free choice of any group for 6 months. This is represented diagramatically below.

People currently on the methadone program, informed consent, random allocation

IH → IH+OM → OM → free choice

IH+OM → OM → IH → free choice

OM → IH → IH+OM → free choice

6 months   6 months   6 months   6 months
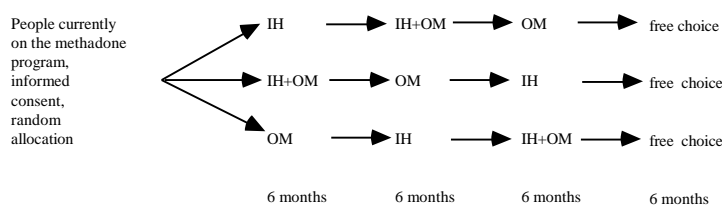
Advantages

- people are less likely to drop–out than in a standard RCT, so that the different options can be more easily compared on a range of outcome measures; thus this is a good design for self–reported outcome measures;
- eliminates reallocation and associated difficulties with the modified RCT;
- this is the best design when amelioration rather than cure is expected and when modest incremental effects are expected.

Disadvantages

- would not allow research questions concerning retention in treatment to be addressed;
- people transferring from a group that has methadone as one of its treatments to a group that does not may have effects caused by withdrawing from methadone (the effects of methadone are more long–lasting than those of heroin);
- the time needed to measure outcomes reliably may become unacceptably long. Outcomes relating to health status, social functioning, criminal behaviour and so on take time to develop, and the time in each treatment group would need to be long enough to clearly relate the outcome to the treatment and to ensure that it was not a carry–over effect from the previous treatment;
- usually cross–over is done with people who are 'blind' to the treatment, but that is clearly not possible here;
- no allowance is made for client choice, as is the case with the modified randomised controlled trial design

Other points

A possible source of bias comes from the likelihood that some (probably many) of the trial participants will enrol primarily to receive heroin. This could mean that those allocated to the heroin option in the first stage may drop out when that stage concludes, whereas those allocated to the non–heroin option first may be motivated to remain on the program because of the incentive of transferring to the heroin option later. This is an argument for including a free choice option as the last stage of the trial. This also highlights the importance of determining (if possible) if dropping–out of the trial was planned by participants (as in the case where trial participants only remain on the trial as long as they are in a particular group) or 'impulsive', in other words motivated by other factors.

7   Outcome Measures

Workshop participants agreed that deciding upon the outcome measures is a prerequisite to making final decisions about trial design. A central issue is whether the measures should focus upon the individual trial participants, be at the community level, or both. The usual approach is to restrict the outcome measures to indicators of the functioning of the individual trial participants, including, where possible, trial drop–outs. Public policy considerations in the case of this study, however, suggest that community–level outcome measures are also relevant. For example, significant decision–makers could well be interested in ascertaining whether or not a trial would cause a reduction in Canberra's crime rate.

The degree to which trial–induced changes among individual participants would be reflected in aggregate community–wide statistics would depend upon the proportion of Canberra's heroin users who participated in the trial and upon the degree to which their behaviour contributed to the aggregate measures. For example, the authorities do not have, and do not expect to have, data on the proportion of Canberra's property crime that is currently committed by illicit drug users. It is impossible to anticipate, then, what community–level changes in property crime could be observed from a successful trial.

Three broad approaches are available for collecting outcome data relating to individual trial participants (including drop–outs), namely:

- self–reported data;
- information on individuals which is gathered from others who are in contact with them; and
- surveillance data (that is, data which are administrative by–products collected by official agencies).

The sorts of information on participants that could be collected from self reports and from people who are in contact with participants include:

- licit and illicit drug use (this could be complemented by objective measures such as urine tests or hair analysis);
- criminal behaviour;
- health;
- social functioning;
- HIV risk behaviour;
- etcetera.

The sorts of information that could be collected through surveillance include:

- arrests, convictions and court dispositions;
- admission to hospital accident and emergency units;
- ambulance attendance at drug overdoses;
- access to medical treatment;
- etcetera.

Clearly, significant privacy considerations are involved here. The matter would need to be approached in a number of ways, including through the informed consent of trial participants, close cooperation between researchers and official agencies and, possibly, through an approach to the Commonwealth Privacy Commissioner. The Australian Capital Territory *Epidemiological Studies (Confidentiality) Act*, which currently covers the Feasibility Research, facilitates the protection of privacy.

A range of other outcome measures could also be relevant. They include the following:

- drop–out or retention. As discussed above, the level of retention may be affected by the design employed, that is, some designs may be more attractive than others, better attuned to the needs of the participants and potential participants. Drop–out or retention is likely to be less of an issue if trial participants are recruited from among the clients of the methadone program (where being allocated to the control group would be the same as the treatment they were currently receiving), than if any dependent heroin users were eligible to participate. In this case though, using drop–out or retention as an outcome measure may have limited value, because the already high retention rate of methadone programs means that only limited scope exists for improvement through the use of the alternatives envisioned for a trial. This has implications in terms of the capacity of the design (especially the number of participants) to identify statistically significant differences in retention levels between the different groups in the trial;
- attracting participants who were not previously in touch with drug–related services. This outcome measure could not be used if eligibility for trial participation was restricted to those currently in a methadone program;
- the processes used in the trial, particularly its service provision component. These processes could also be described and assessed as outcome measures. This reflects the fact that, for many of the clients and providers of drug treatment services, the quality of the interpersonal interaction in the service is significant in its own right, as well as being a factor influencing retention rates.

It was also suggested that the trial could have both negotiable and non–negotiable outcome measures. In other words, to be eligible to enter the trial, potential participants would have to agree to provide certain information, and agree that the researchers could obtain specified information from third parties. The latter could include surveillance data from official sources. Accessing other specified outcome indicators could be a matter for negotiation. Obtaining information from friends and relatives could fall into this category.

8   Non–Experimental Designs

Although most of the discussion at the workshop focused upon the experimental designs discussed above, some consideration was also given to non–experimental approaches. As touched upon above, the basic concept put forward was whether individual heroin users in Canberra should be the unit of analysis in the research, or whether the focus would be better placed on the Canberra community as an entity, or on Canberra heroin users as a community. (It was pointed out that there is no real 'community' of heroin users in Canberra. Rather, there are a number of networks of users and the members of different networks often do not interact intensively.)

Discussion revolved around the need to clarify the research question. For some members of the public and decision makers, the key question could be something like 'what is the likely impact on the Canberra heroin–using community of making the drug available in a manner different from the way it is at present?'. The outcome measures that would flow from such an approach would be primarily at the community–wide level (as discussed above), rather than relating to the functioning of individual trial participants. This is a complex research question, one that does not lend itself nicely to experimental research designs. (Adequately controlled comparisons of different regimes of heroin availability in different Australian cities, an experimental approach using the community as the level of analysis, was not considered feasible.) Focusing on the community rather than individual heroin users is also less intrusive and recognises the potential vulnerability of individual users.

The quasi–experimental design of the 'n=1' variety was mentioned. In this approach, the Canberra heroin–using community (or the total Canberra community) would be the unit of analysis. Measures would be made of the nature of the community at Time 1, heroin would be made available in a controlled manner for a period and the nature of the community at Time 2 would be assessed and differences between Time 1 and Time 2 observed. This would be a prospective design with an historical control. The weakness of this design, it was agreed, is the multiple confounders of the measures: many changes could occur in the community between Time 1 and Time 2, resulting from diverse influences. It would be difficult or impossible to assess the changes caused by the changed availability of heroin. Most workshop participants therefore favoured one of the experimental approaches discussed earlier.

Nevertheless, the discussion about non–experimental approaches was valuable in clarifying the limitations of the experimental approaches. It threw into relief the limits of generalisability of an experimentally based trial, the types of research questions that such a trial could ask and the units of analysis for such a trial. The discussion also highlighted the political aspects of trial design and evaluation and the need, at a minimum, to document political constraints. There was some discussion about the extent to which trial design and evaluation should be influenced by political "realities".

9   Conclusion

There was a level of agreement at the workshop that the unit of analysis should be heroin users willing to enter treatment, rather than all heroin users, and that the primary research question should be concerned with the quality of life of the trial participants. As discussed earlier, it was concluded that the results of an experimentally based trial would be generalisable only to treatment programs that were similar to the

trial. An experimentally based trial would provide no information on other issues such as the likely impact of controlled availability of heroin on the community generally, though such information could possibly be gathered through research external to the trial.

Final decisions about trial design and evaluation cannot be made until research which is currently underway is completed. This research allows for a more refined estimation of potential numbers of trial participants and likely levels of health problems, HIV risk behaviours, criminal behaviours, illicit drug use and social problems associated with heroin use among potential participants. This will determine whether or not such variables can be used as outcome measures.

10   Workshop Participants

Gabriele Bammer, NCEPH

Jeanne Daly, Department of Sociology, La Trobe University

Richard Jarrett, Department of Statistics, University of Adelaide

Mike Lane, NCEPH

David McDonald, AIC

David Newell, Department of Community Medicine, Westmead Hospital

Aileen Plant, NCEPH

Lawrence Sherman, Crime Control Institute, Washington DC

Beverly Sibthorpe, NCEPH

Patty Solomon, Department of Statistics, University of Adelaide

Sue Wilson, Centre for Mathematics and Its Applications, Australian National University

Alex Wodak, Alcohol and Drug Service, St Vincent's Hospital

References: provided as background reading

Bammer, G. (1993), 'Should the controlled provision of heroin be a treatment option? Australian feasibility considerations', *Addiction* 88: 467–475.

Bammer, G., Douglas, R. & Dance, P. (1991), 'Evaluation by a randomised controlled trial', in National Centre for Epidemiology and Population Health, *Feasibility Research into the Controlled Availability of Opioids Stage 1: Volume 2 Background Papers*, 279–286.

Jarrett, R.G. & Solomon, P.J. (1994), 'An evaluation of possible designs for a heroin trial', see this working paper.

McDonald, D.N., Bammer, G., Legge, D.G. & Sibthorpe, B.M. (1994), 'Service provision considerations for the evaluation of a heroin trial. A discussion paper', see this working paper.

Ostini, R., Bammer G., Dance, P.R. & Goodin, R.E. (1993), 'The ethics of experimental heroin maintenance', *Journal of Medical Ethics*, 19: 175–182.

References: not provided as background reading

Sherman, L.W. (1992), *Policing Domestic Violence: Experiments and Dilemmas*, Free Press: New York.

AN EVALUATION OF POSSIBLE DESIGNS FOR A
HEROIN TRIAL
(DRAFT)

R.G. Jarrett & P.J. Solomon
Department of Statistics
University of Adelaide

16 August 1993

Abstract

This paper considers several alternative trial designs in the case where some participants elect not to have certain treatments. It is shown that this necessarily leads to a loss of power, but that substantial improvements can be made over the 'intention to treat' analysis. Two main designs are considered, with one showing a clear advantage over the other.

1 Introduction

A trial to provide heroin in a controlled manner to dependent users must be rigorously designed, conducted and evaluated. During the Stage 1 feasibility investigations, it was proposed that this could best be done in by providing heroin in a randomised controlled trial (Bammer et al., 1991). However this methodology has important limitations (Bammer, 1993).

In recent years there has been growing resistance by participants and, often, service providers to randomised controlled trials. This may manifest in refusal by service providers to allocate participants, pooling of medications by participants and high drop-out rates. A randomised controlled trial comparing heroin and methadone has only been tried once before. At that time injectable heroin was the usual treatment and oral methadone the new alternative. There was a high drop-out rate from the oral methadone group. Twelve percent of participants allocated to oral methadone left the trial immediately. By three months less than 50% were still in the trial and by 12 months this had fallen to 29%. At 12 months 74% of those in the heroin arm were still in the trial (Hartnoll et al., 1980).

More recently, traditional designs and conduct of randomised controlled trials have been challenged by drug trials for the treatment of people with HIV disease. The major issues and difficulties in this context, and some suggestions for overcoming them, are discussed in Byar et al (1990) and in Green et al (1990). The emphasis in rethinking is on making such trials flexible and acceptable to potential participants.

One flexible approach that has been tried is the use of *subtrials.* In the AIDS trial 'Alpha', patients intolerant of AZT were offered a choice of subtrials: patient choosing subtrial A were randomised onto high DDI vs low DDI vs placebo; patients choosing subtrial B were randomised onto high DDI vs low DDI. However, the randomisation scheme did not work very well, in that very few participants opted for the subtrial containing the placebo group and this arm of the trial was wound up early (personal communication with Professor Peter Armitage). The subtrial idea also comes in the cardiovascular trial ISIS-3 (1992), where patients for whom physicians thought fibrinolytic therapy, was required were randomised to streptkinase, tPA or APSAC, whereas those for whom the indication was 'uncertain' were randomised to one of these three groups or to 'open control', but no placebo.

The use of subtrials in these contexts has motivated us to consider the prospect of their use for the heroin trial. This report details several options which are available for the conduct of the heroin trial. These options are based around the idea that it is expected that some participants will not agree to certain treatments and that it is therefore important to allow a certain element of choice.

We will consider three different treatments to which participants may be allocated. These will be referred to as H (heroin only); H+M (heroin and methadone combined) and M (methadone only).

In looking at the comparison of treatments, we will suppose that each observation has a variance σ^2 and that there are N participants altogether. At this stage, we do not need to discuss what measurements are actually being made on each participant. The figures quoted will then be the standard deviations of treatment estimates, ignoring the factor $\sigma/\sqrt{N}$. We will want this number to be as small as possible.

The most important treatment estimate to consider is the difference between the Methadone group (M) and the Heroin group (H). In all the models considered, the number of participants and the method of allocation will be the same for both H and H+M, so that the comparison between the H+M and the M groups will have the same accuracy. We might expect that the comparison of the H and H+M groups will be more accurate in general because of the way participants are allocated to those groups.

The baseline figure would be if we simply could randomly assign all N participants into three equal size groups. The result would be a variance for the treatment differences of $6\sigma^2/N$, so the standard deviation figure ignoring the factor $\sigma/\sqrt{N}$ is $\sqrt{6} = 2.449$. We will use the symbol R to denote the figure we get from various designs, remembering that $R = 2.449$ is the best we can hope for.

The analysis at this stage will make several simplifying assumptions which will be outlined in the course of this paper. A more thorough study will have to address these issues and would be necessary before the results can be published. We do not, however, believe that these assumptions will materially alter the results contained in this report.

2 Design I

If the participant is suitable for the trial, we would first ask whether the participant is prepared to accept a treatment which does not involve the use of heroin. If the answer is NO, they go into Group I; if YES, they go into Group II. (It is possible that there will be a Group III consisting of those who will only accept NON-heroin treatments or in fact the H group alone.) The treatment allocation can then be pictured as:

	H	H+M	M
Group I	A	B	
Group II		C	D

and we might put $N/4$ of the N participants into each of the 4 possibilities. Since there is randomization into the two treatment choices within each of the groups, the treatment comparisons within Group I and within Group II are valid. However, comparisons across the groups cannot be justified on a randomization argument because of the self-selection of participants into the two groups. It is possible to make a comparison between the two H+M groups, since they have the same treatment regime, so that any difference can be ascribed to the bias caused by the element of self-assignment. The difference between these two groups can be regarded as the bias which needs to be corrected. If we let the baseline reading be a, the difference between the H and the H+M groups be b. the effect of self-selection be c, and the difference between the H and M groups be d, the expected values in the four cells are:

	H	H+M	M
Group I	a	a+b	
Group II		a+b+c	a+d+c

Labelling the four treatment regimes above *A, B, C* and *D*, we can say that the estimate of the difference *d* between Heroin and Methadone would be estimated by,

$$(A - B) + (C - D),$$

where the difference *(B - C)* corrects for the bias, just as a covariate would. Each mean has a variance of $4s^2/N$, so the variance of this combination of means is given by $16s^2/N$. Thus, the ideal figure for the standard error of the estimate of 2.449 mentioned above has been increased to 4.00.

Note that this value would be increased if

- the ratio of sample sizes *A : B* or *C : D* differed from 1: 1;
- the choice for insisting on heroin in the treatment differed from 50%.

We need to explore this further as there may only be 10%, say, who insist on having heroin in the treatment and we would then only get 5% of all the patients having the first treatment regime.

Suppose that a proportion *p* of the *N* insist on having heroin (i.e. being in Group I). The optimal allocation to minimize the variance above is to have proportions $p/2$ in each of *A* and *B*, and $(1 - p)/2$ in each of *C* and *D*. The variance above then becomes $4\sigma^2/\{Np(1 - P)\}$, which gives a value of *R* which is very close to 4 for reasonable departures of *p* from 50%. For example, at $p = 0.20$, we get 5.00, and at $p = 0.10$ it increases to 6.67.

If *p* is large, we have an unavoidable problem because we simply cannot get many, patients into the M group no matter what we do. However, if *p* is small, the current design allocates very few to the H group, and this *is* avoidable. One way to do this is to allocate the Group II participants to all three of the treatments rather than just to the H+M and M treatments and this is the participant of the next section.

3 Design II

3.1 The Model

We will again use as our benchmark the comparison of the H group with the M group. In this case, suppose that a proportion *p* opt for heroin only (i.e. H or H+M) and the remainder don't mind. We now have to consider (i) the subdivision of each of Group I and II into the treatment regimes, and (ii) the estimation of the difference between the estimates of H and the estimate of M.

The layout now has 5 possible treatment groups, the 4 given in the previous section plus a fifth one that we will label *E*:

	H	H+M	M
Group I	A	B	
Group II	E	C	D

Again, we use as our model an additive model whereby there is a "block" difference between Group I and Group II, plus treatment effects corresponding to the three treatment regimes. Our model then is that the fitted values look like:

	H	H+M	M
Group I	a	a+b	
Group II	a+c	a+b+c	a+d+c

and we are interested in the estimate of d, the difference between Heroin and Methadone, unencumbered by the difference c between the groups.

Clearly, one estimate of this is just $D - E$ and this is valid due to the randomization. If, however, we want to avail ourselves of the information in Group I, we can estimate b and c and hence arrive at a solution.

For simplicity, we will assume that we divide Group I equally between the two treatments, and for Group II we choose a proportion (1 - f) to get Methadone only, and divide the rest equally between the other two groups. The proportions are then:

	H	H+M	M
Group I	p/2	p/2	
Group II	(1-p)f/2	(1-p)f/2	(1-p)(1-f)

and we can now work out the estimates and the variances.

3.2 The Analysis

As a starting point, we note that

$$\text{Var}\,(D-E) = \frac{s^2}{N}\,\frac{(2-f)}{(1-p)f(1-f)}\;,$$

but this ignores the information in groups A and B. If p were zero, this would of course do very well, and we could get our figure slightly below 2.449 by a suitable choice of f. There are 5 group means and 4 parameters to estimate here, so we can find a more accurate estimate by performing a least squares analysis of the model given above. Details of these calculations will be given in an Appendix. After some calculation, we arrive at the formula for the variance as

$$\text{Var}\,(d) = \frac{s^2}{N(1-p)f(1-f)}\; \frac{\{p+(1-p)f(2-f)\}}{\{p+(1-p)f\}}\,.$$

We note the similarity to the previous variance, particularly that this variance is necessarily smaller unless $p = 0$.

Apart from the factor σ^2/N which we ignore for the moment, we can calculate this for any given value of p and f. The results shown in Table 1, represent R; that is, we drop the factor σ^2/N from the formula and then take the square root. Clearly, by symmetry, the comparison of M with H+M would have the same variance. The accompanying Figure 1 shows the plot of these values, with increasing values of p from 0 to 0.9 given by the different curves, starting at the bottom of the graph with $p = 0$.

The conclusion is that the standard deviation is heavily influenced by p, high values of p giving much higher standard errors for the treatment comparisons. However, for a given value of p, the standard deviation is very robust to different values of f, with any value between 0.3 and 0.8 being acceptable. The best value for f is around 0.60 when p is small, but is closer to 0.50 as p increases. The increase in the standard deviation as p increases is small, being only a 10% increase when p is as high as 30%.

Table 1: standard deviation of estimate (H-M) given the proportion "p" insisting on heroin and the proportion "f" of the remainder who are divided equally into H and H+M.

Proportion insisting on heroin

		0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
	0.05	6.407	5.503	5.521	5.750	6.117	6.634	7.365	8.460	10.320	14.548
	0.10	4.595	4.196	4.179	4.310	4.549	4.903	5.417	6.197	7.535	10.593
	0.15	3.809	3.601	3.596	3.698	3.888	4.174	4.596	5.243	6.358	8.918
	0.20	3.354	3.243	3.254	3.347	3.513	3.764	4.135	4.706	5.696	7.974
	0.25	3.055	3.000	3.028	3.118	3.272	3.502	3.842	4.367	5.277	7.377
	0.30	2.845	2.827	2.868	2.960	3.108	3.326	3.646	4.140	4.997	6.978
	0.35	2.693	2.701	2.753	2.849	2.994	3.205	3.513	3.987	4.809	6.710
	0.40	2.582	2.608	2.670	2.771	2.917	3.124	3.425	3.887	4.687	6.537
f	0.45	2.503	2.544	2.615	2.720	2.868	3.076	3.374	3.830	4.618	6.439
	0.50	2.449	2.503	2.582	2.693	2.845	3.055	3.354	3.809	4.595	6.407
	0.55	2.420	2.484	2.572	2.689	2.847	3.061	3.365	3.824	4.615	6.438
	0.60	2.415	2.488	2.584	2.709	2.874	3.096	3.407	3.876	4.682	6.535
	0.65	2.436	2.519	2.624	2.758	2.931	3.163	3.486	3.971	4.801	6.708
	0.70	2.488	2.581	2.696	2.841	3.026	3.271	3.611	4.120	4.987	6.975
	0.75	2.582	2.686	2.814	2.972	3.173	3.436	3.801	4.343	5.265	7.373
	0.80	2.739	2.857	3.000	3.177	3.399	3.689	4.088	4.679	5.683	7.970
	0.85	3.003	3.142	3.308	3.510	3.764	4.095	4.547	5.214	6.344	8.913
	0.90	3.496	3.667	3.870	4.117	4.425	4.824	5.368	6.170	7.522	10.589
	0.95	4.702	4.944	5.230	5.578	6.010	6.567	7.325	8.437	10.309	14.544

3.3 An Example

One way to use this table is to consider how much the sample size would have to increase to produce the figure 2.449. The answer to this is the square of the ratio of the two numbers. Thus, if $p = 0.4$, 40% of the participants decline the Methadone only treatment and our best value of f is 0.50. The proportions in the groups are then:

	H	H+M	M
Group I	0.2	0.2	
Group II	0.15	0.15	0.3

We then obtain $R = 2.845$. To achieve the same accuracy from the experiment as we would have got with $p = 0$, we need to increase our sample size from N to $N(2.845/2.449)^2 = 1.35N$. Thus, we would need 35% more participants to achieve the same accuracy as the $p = 0$ case. Of course, if we assumed $p = 0$ by not asking, the variance in the M group would increase because some participants would opt out and seek other sources of heroin.

4   Design III

The alternative is to consider properly randomised designs in which we then give some participants choice as to whether they opt out of a particular treatment. The scenario here is that we explain to participants about the clinical trial and the three possible treatments. They are told that they will be allocated to one of the three treatments but that, if they are allocated to methadone only (M), they will have the option to refuse and be switched to one of the other treatments. If they choose this option, they would be randomized to either H or H+M.

4.1 The Layout

This produces a design very similar to the second option above. We can produce the same layout into 5 treatment classes, but Group I now corresponds only to those people who were first allocated to M and then revealed that they would not accept M. Others who not accept M are now allocated directly to H or H+M and never asked about their preferences. In the table below, E and C represent the groups initially randomized into H and H+M. The group assigned to M are then given the choice of opting out; we assume a proportion p do so, and are randomly assigned (in equal proportions) to A and B; those who do not opt out remain as Group D:

	H	H+M	M
Group I	A	B	
Group II	E	C	D

The decision we have to make is what proportions are allocated to the 3 groups initially. We will suppose that a proportion $f/2$ are allocated to each of E and C, and the remaining proportion $(1-f)$ are allocated to Methadone. This leads to the following proportions in the groups:

	H	H+M	M
Group I	p(1-f)/2	p(1-f)/2	
Group II	f/2	f/2	(1-p)(1-f)

There are two ways to analyse this trial. The first is to fit a model similar to that in Design II, where we fit a linear model to the 5 groups formed, assuming as before that there is a difference between the heroin based groups according to whether they accept or decline to be in the Methadone group. The second method is to do it according to the "intention to treat", so that we just consider the three groups formed by the initial randomization. It follows that the Methadone group will in fact be a mixture of results from all three treatments, with the consequent loss of power, particularly when p, the proportion opting out of Methadone, is large.

4.2 The Model

As a first step, we consider the same model as for the previous design, namely that
• those in the H+M group have a mean which is higher by b than those in the H+M group,
• those who would be prepared to be in the M group have a mean higher by c than those who would (if asked) have opted out of the M only treatment, and
• those in the M group have a mean higher by d than those who are in the H group and who would have agreed to be in the M group (if asked).

Again, our primary interest will centre on estimating the parameter d. To obtain the expected values in the groups, we note that both E and C are in fact a mixture of two sets of participants, namely a proportion p who would have declined to be in the M group had they been offered it, and a proportion $(1-p)$ who would have accepted the M group if offered. They were not, given the choice. The reality is that the number who would have made each choice has a binomial distribution with these proportions. In our calculations, we assume it is exactly a proportion p. Then the mean for Group E, for example is given by

$$p \times a + (1 - p) \times (a + c) = a + c(1 - p),$$

while for Group C it is $a + b + c(1 - p)$. Thus, the table of means for the 5 groups is

	H	H+M	M
Group I	a	a+b	
Group II	a+c(l-p)	a+b+c(l-p)	a+c+d

with the proportions given above. The mixtures in Groups C and E imply an increase in variance. For Group E, a proportion p have mean a and a proportion $(1 - p)$ have mean $a + c$. This implies that the variance of a single observation is increased from the usual σ^2 to $\sigma^2 + c^2 p(1 - p)$.

The same variance applies to a single observation from group c. Since the distribution of values from these groups is a mixture of two normal distributions, the variance does not decrease in the usual way as we take more observations and take a mean. In very large samples, for example, we will have a concentration of values about the two separate values corresponding to the two types of observations making up the group. Hence, when there are n participants with the proportions in each group as above, the variances of the group means will be

$$\mathrm{Var}(C) = \mathrm{Var}(E) = 2\sigma^2 / (N\,f) + c^2 p(1 - p).$$

We do not pursue this here, and will work on the basis that c is small while arguing that it should be included in the estimation procedure.

4.3 The Analysis

Now we can argue here for two approximations. The first is that, unless we ask who would or who wouldn't have accepted the M group allocation, the sample variances calculated in each group don't contain much information about the value of c. We would need large sample sizes for this to be useful, as such information would rely on the ratio of two sample variances , which are generally highly variable. The second is that we will assume for the present that c is small relative to the other effects and hence that, unless N is large, the effect of c on the variance of the group mean is marginal. Suppose for example that $p = f = 0.5$ and that $c = 0.5\sigma$. Then the variance of the group mean for E is increased by a factor $1 + N/64$. If we run a trial with several hundred people, this will severely reduce the accuracy of these two group means and will mean the results that follow would need to be reworked. For the moment we ignore this effect.

With these two simplifying assumptions, we may proceed to the least squares analysis. The variances of the group means are σ^2 /N divided by, the proportions allocated to each of the five groups. We then find that the variance of d is given by

$$\text{Var}\,(d) = \frac{s^2}{N(1-p)f(1-f)}\ \frac{[p+(1-p)f\{p+(1-p)(2-f)\}]}{\{p+(1-p)f\}(1-p)}\ ,$$

which can be shown, for given p and f to be necessarily larger than the formula given for Design II. This does not necessarily matter because we may well choose to use a different f for this design. However, it does imply that, for any given p, we can choose an f which will give a lower variance. Table 2 shows the values for a range of p and f, and Figure 2 is a graph of the same values with p increasing from 0 in the bottom curve to 0.9 for the top curve.

4.4 An Example

Consider the same example. If we think that p is about 0.4, we would use $f = 0.5$ anyway. Thus, initially, we allocate 25% of the participants to each of H and H+M, and 50% to the M group. If 40% of the M group opt out and get allocated to the other groups, the final allocation is:

	H	H+M	M
Group I	0.15	0.15	
Group II	0.25	0.25	0.20

The value of R is 3.541, so the sample size has to increase to 2.09N, even larger than before.

TABLE 2: Standard deviation of estimate (H-M) in Design III, given the proportion "p" insisting on heroin and the proportion "f" of the remainder who are divided equally into H and H+M (Least Squares model).

Proportion "p" insisting on heroin

f	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.05	6.407	5.735	6.088	6.779	7.798	9.280	11.541	15.340	22.968	45.895
0.10	4.595	4.357	4.575	5.038	5.748	6.802	8.427	11.173	16.703	33.350
0.15	3.809	3.733	3.922	4.298	4.881	5.754	7.109	9.407	14.046	28.025
0.20	3.354	3.360	3.542	3.877	4.391	5.164	6.367	8.412	12.548	25.022
0.25	3.055	3.108	3.291	3.604	4.078	4.789	5.896	7.781	11.598	23.117
0.30	2.845	2.929	3.117	3.417	3.867	4.537	5.581	7.360	10.964	21.846
0.35	2.693	2.799	2.992	3.287	3.721	4.366	5.369	7.077	10.538	20.991
0.40	2.582	2.705	2.904	3.198	3.624	4.254	5.230	6.893	10.262	20.438
0.45	2.503	2.640	2.846	3.142	3.566	4.188	5.151	6.789	10.106	20.127
0.50	2.449	2.599	2.814	3.115	3.541	4.163	5.123	6.754	10.055	20.026
0.55	2.420	2.582	2.806	3.115	3.549	4.178	5.145	6.786	10.105	20.127
0.60	2.415	2.590	2.825	3.145	3.590	4.233	5.218	6.887	10.259	20.438
0.65	2.436	2.624	2.874	3.209	3.671	4.335	5.351	7.068	10.534	20.990
0.70	2.488	2.692	2.959	3.314	3.801	4.497	5.559	7.349	10.960	21.845
0.75	2.582	2.806	3.096	3.478	3.999	4.741	5.869	7.768	11.592	23.116
0.80	2.739	2.989	3.310	3.731	4.301	5.110	6.336	8.397	12.542	25.020
0.85	3.003	3.292	3.659	4.138	4.784	5.697	7.077	9.391	14.040	28.024
0.90	3.496	3.849	4.295	4.874	5.650	6.745	8.396	11.157	16.697	33.348
0.95	4.702	5.200	5.825	6.633	7.714	9.232	11.515	15.328	22.964	45.894

4.5 Intention to Treat

An alternative analysis here is to ignore the randomisation at the second stage and simply to analyse the data on the basis of the original randomization; that is, on the basis of the intention to treat. This will necessarily throw away known information since we do know what treatments the participants eventually had, whereas the intention to treat usually works on the basis that we don't really know who used what in the M group. We would therefore expect the variances to be larger.

To do this, we need to combine groups A, B and D and consider them as one group. This leads to the following expected values:

	E	C	F={A,B,D}
Mean	a+c(1-p)	a+b+c(1-p)	a+c(1-p)+pb/2+(1-p)d
Propn	f/2	f/2	(1-f)

From this we see that we cannot distinguish any more between a and c, and there are three means which essentially depend on only three parameters. Our solutions are therefore given by

$$a+c(1-p) = E$$
$$b = C-E$$
$$d = \{F-pC/2-(1-p/2)E\}/(1-p)$$

where C, E and F represent the mean values for each group. Note that, as mentioned earlier, the mixing of groups with different means leads to additional variation for the mean that does not decrease as N gets larger. If, however, we again ignore this portion of the variation, we can evaluate the variance of this linear combination and obtain:

$$\text{Var}(d) = \frac{s^2[f + (1-f)\{1 + (1-p)^2\}]}{(1-p)^2 f(1-f)} \ ,$$

which again can be shown to be larger than the previous formulae. The results for this are in Table 3, and are depicted graphically in Figure 3.

TABLE 3: Standard deviation of estimate (H-M) in Design III, given the proportion "p" insisting on heroin and the proportion "f" of the remainder who are divided equally into H and H+M (Intention to Treat model) .

Proportion "p" insisting on heroin

f	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.05	6.407	6.782	7.273	7.935	8.859	10-208	12.312	15.935	23.373	46.101
0.10	4.595	4.870	5.231	5.716	6.393	7.379	8.913	11.552	16.964	33.483
0.15	3.809	4.043	4.350	4.762	5.334	6.168	7.462	9.686	14.239	28.124
0.20	3.354	3.566	3.843	4.214	4.729	5.477	6.638	8.628	12.698	25.100
0.25	3.055	3.253	3.512	3.858	4.338	5.033	6.110	7.954	11.719	23.180
0.30	2.845	3.035	3.282	3.613	4.070	4.731	5.753	7.500	11.063	21.898
0.35	2.693	2.878	3.119	3.439	3.882	4.521	5.507	7.190	10.618	21.034
0.40	2.582	2.765	3.002	3.317	3.752	4.378	5.342	6.985	10.328	20.474
0.45	2.503	2.685	2.922	3.235	3.667	4.288	5.242	6.864	10.160	20.156
0.50	2.449	2.634	2.872	3.188	3.621	4.243	5.196	6.815	10.100	20.050
0.55	2.420	2.609	2.852	3.172	3.611	4.240	5.203	6.835	10.140	20.146
0.60	2.415	2.610	2.860	3.189	3.639	4.282	5.264	6.926	10.288	20.453
0.65	2.436	2.639	2.899	3.242	3.708	4.373	5.386	7.098	10.556	21.002
0.70	2.488	2.703	2.978	3.339	3.828	4.525	5.585	7.371	10.976	21.854
0.75	2.582	2.814	3.109	3.495	4.018	4.761	5.888	7.784	11.605	23.123
0.80	2.739	2.994	3.319	3.742	4.314	5.123	6.349	8.408	12.550	5.025
0.85	3.003	3.295	3.665	4.145	4.792	5.705	7.085	9.398	14.045	28.027
0.90	3.496	3.851	4.298	4.877	5.655	6.749.	8.400	11.161	16.700	33.350
0.95	4.702	5.200	5.826	6.635	7.716	9.234	11.517	15.329	22.965	45.895

5   Conclusions

Figure 4 shows the actual randomization scheme used for Design II and Design III (indicated there as Solution 1 and Solution 2) and makes clear the differences between them.

The conclusions reached are the following:

• It is clear that as soon as we have some participants declining particular treatments, we have an increased standard deviation for comparisons; this is the same as a loss of power - that power can only be regained by increasing the sample size.

• When participants decline particular treatments, we cannot assume that their final outcome measurements will follow the same pattern as those who do not decline those treatments. This study supposes that there is a shift in the mean between those two groups.

• The loss of power occurs most rapidly using the Intention to Treat model for Design III. This is because the analysis does not utilise known information about which participants chose which treatments.

• Design II and Design III are essentially giving similar results. In practice we will not know what the value of p is prior to starting the experiment, so it would be wise to choose a value of f which is conservative in the sense of affording protection against a larger rather than a smaller value of p. A value close to 0.50 appears to be the best value to take.

• The analysis shows a preference for Design II as p gets larger. Once p is about 0.4, Design III needs 50% more participants (compare the figures $2.09N$ and $1.35N$ above) to achieve the same power.

• The question which must be considered is whether the linear model is suitable (on some scale) for the outcome measures envisaged for this experiment. If the outcome measure is related to the time taken to some event, then the linear model might well apply, as in a Cox model to the log hazard.

• The current analysis is incomplete for several reasons. The crucial assumption is that we have ignored the effect on the variance of group means caused by the fact that they are often mixtures of observations from different groups. This will inflate the estimates of variance for Design III, which is where this happens, for both the ordinary analysis and the intention to treat analysis. More work needs to be done on this aspect. This suggests that Design II is a preferable scenario. Design II is essentially running two separate trials but attempting to combine the result afterwards.

6 References

BAMMER, G. (1993) Should the controlled provision of heroin be a treatment option? Australian feasibility considerations. Addiction, 88, pp. 467-475.

BAMMER. G., DOUGLAS, R. & DANCE, P. (1991) Evaluation by a randomised controlled trial, in: NATIONAL CENTRE FOR EPIDEMIOLOGY AND POPULATION HEALTH. Feasibility Research into the Controlled Availability of Opioids Volume 2. Background Papers, pp 279-286 (Canberra, ACT, NCEPH).

BYAR, D.P. AND 21 OTHERS (1990) Design considerations for AIDS trials, New England Journal of Medicine, 323, pp. 1343-1348.

GREEN, S.B., ELLENBERG, S. S., FINKELSTEIN, D., FORSYTHE, A. B.; FREEDMAN, L. S., FREEMAN, K., LEFKOPOULOU, M., SCHOENFELD, D. & SMITH, R. P. (1990) Issues in the design of drug trials for AIDS, Controlled Clinical Trials, 11, pp. 80-87.

HARTNOLL, R. L., MITCHESON, M. C., BATTERSBY, A., BROWN, G., ELLIS, M., FLEMIING, P. &, HEDLEY, N. (1980) Evaluation of heroin maintenance in controlled trial, Archives of General Psychiatry, 37, pp. 877-884.

ISIS-3 COLLABORATIVE GROUP. (1992) ISIS-3: a randomised comparison of streptokinase vs tissue plasminogen activator vs anistreplase and of aspirin plus heparin vs aspirin alone among 41,299 cases of suspected acute myocardial infarction, The Lancet, 339, pp. 753-770.
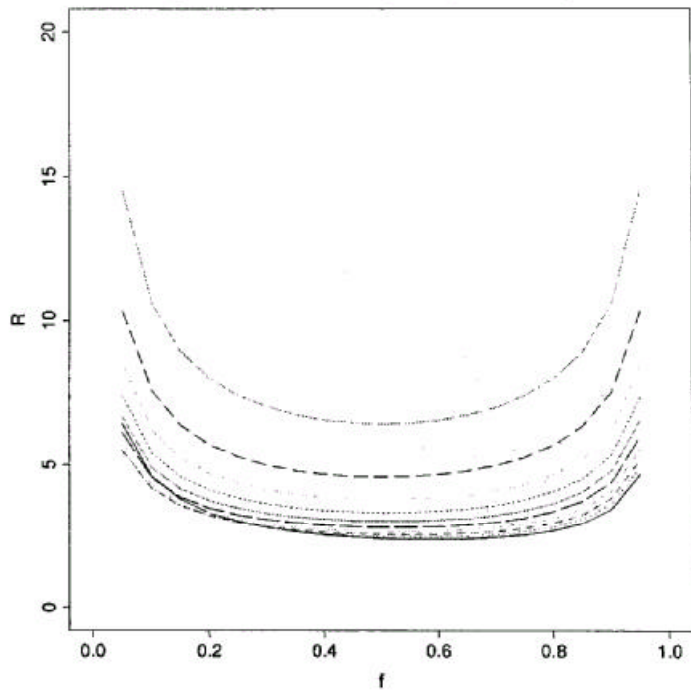
Figure 1: Plot of SD(H-M) for Design II

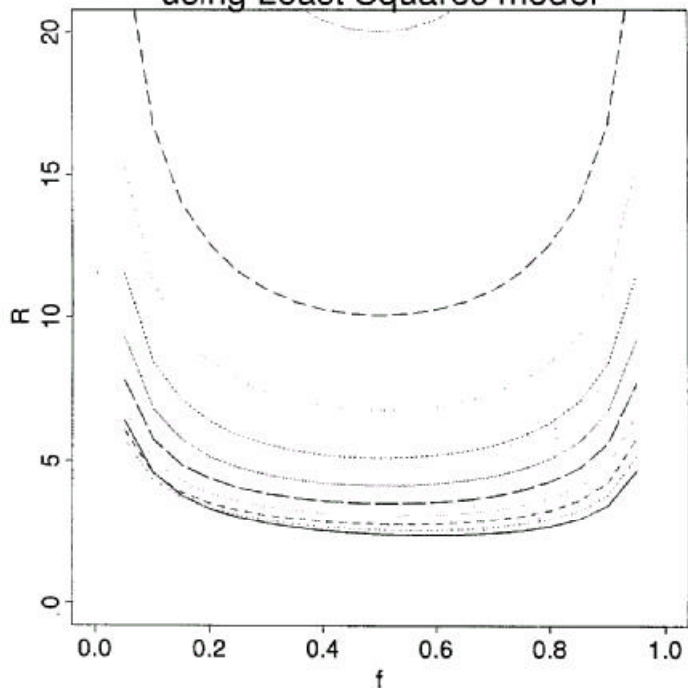Figure 2: Plot of SD(H-M) for Design III, using Least Squares model

Figure 3: Plot of SD(H-M) for Design III, using Intention to Treat model
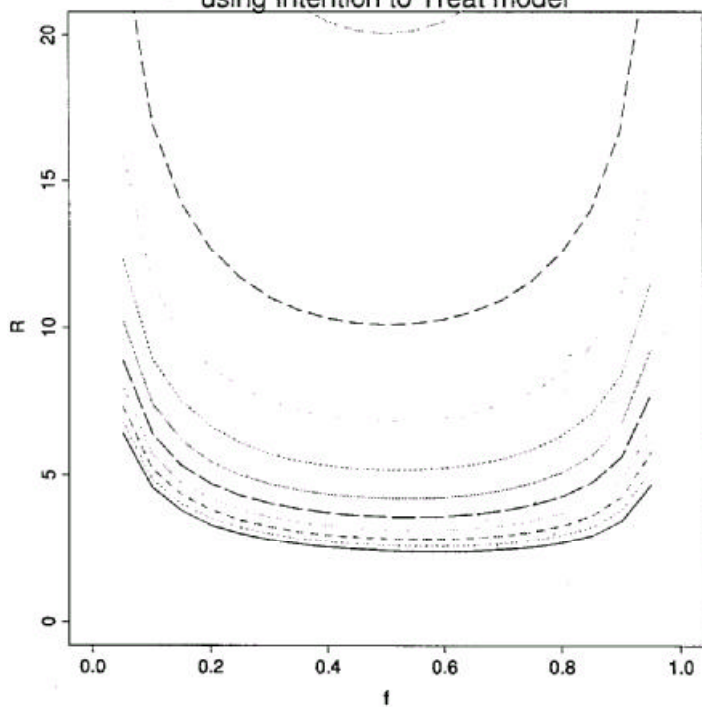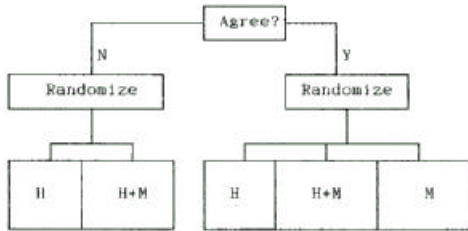
Figure 4

SOME POSSIBILITIES FOR HEROIN TRIAL

SCENARIO: Three possible treatments: Heroin, Heroin + Methadone, Methadone only.
PROBLEM: Some potential participants will not accept Methadone only.

SOLUTION 1: Find out which first, then randomize.
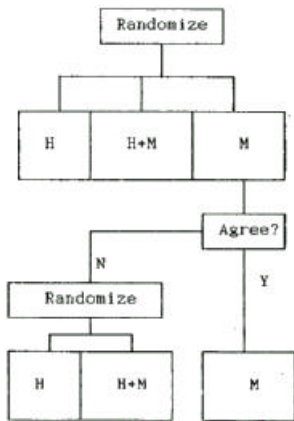

SOLUTION 2: Randomize first, then ask.
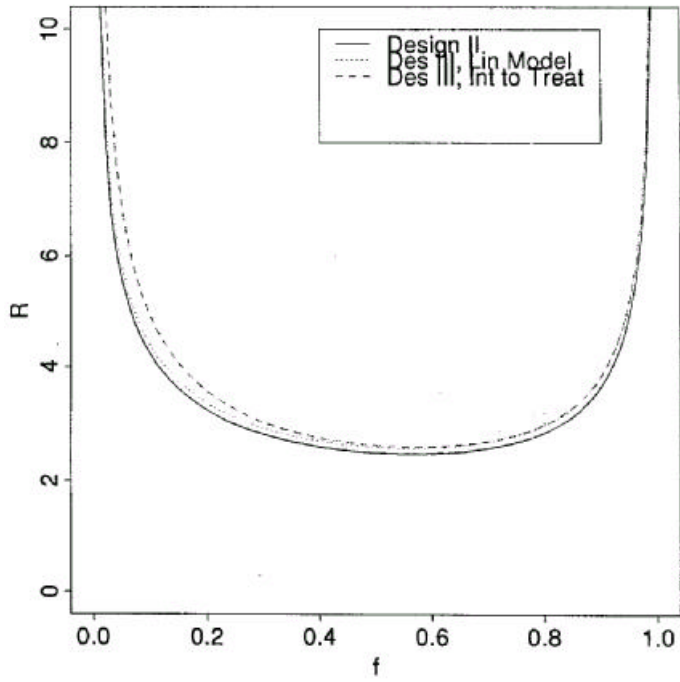   If necessary, re-randomize.

Figure 5: Plot of SD(H-M) when p=0.1

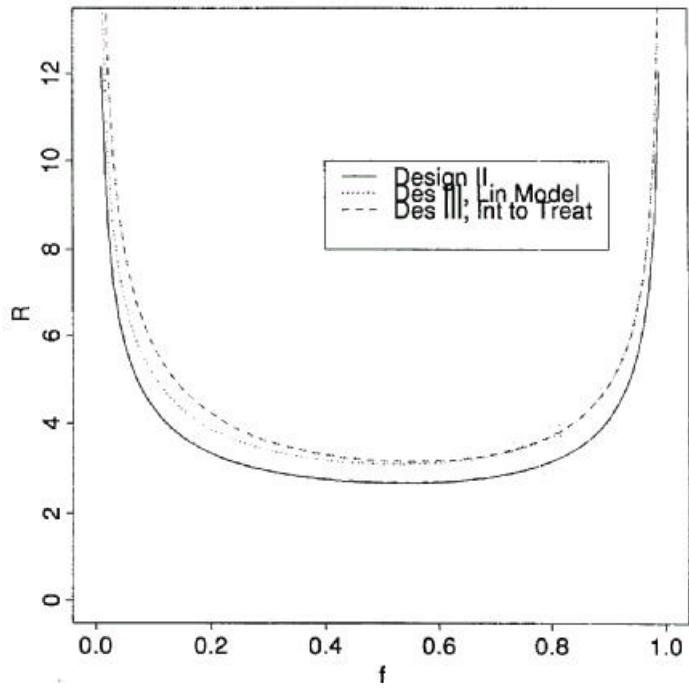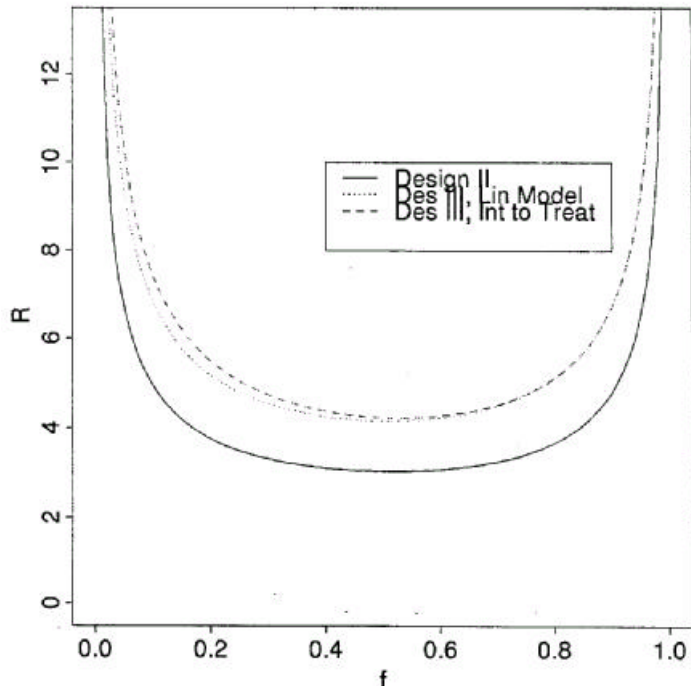Figure 6: Plot of SD(H-M) when p=0.3

Figure 7: Plot of SD(H-M) when p=0.5

SERVICE PROVISION CONSIDERATIONS FOR THE EVALUATION OF A 'HEROIN TRIAL' A DISCUSSION PAPER

David McDonald[#], Gabriele Bammer[*], David G. Legge[*], Beverly M. Sibthorpe[*]

[*] National Centre for Epidemiology and Population Health
[#] Australian Institute of Criminology

1   Introduction

This discussion paper raises for consideration issues concerned with service provision within a proposed trial of controlled availability of heroin. The thesis of this paper is that while we recognise the strengths of a randomised controlled trial in answering certain types of research questions, there are service provision concerns of trial participants (both staff and heroin users) that may be inconsistent with standard RCT designs. These concerns relate to the effectiveness of service provision. There is widespread concern about the effectiveness of health care services generally, both in relation to successful outcomes and cost effectiveness. With regard to service provision for drug users, there is generally considered to be considerable room for improvement in both prevention and treatment strategies. We reflect here on the principles we consider would underlie the provision of an effective service in the context of the controlled availability of heroin. They constitute a health development approach.

What are the objectives of the research? The research question can be conceived as something like 'Does the controlled availability of heroin result in less harm than the current uncontrolled availability?'. We prefer, however, a formulation more like 'What are the conditions under which the controlled availability of heroin will result in less harm than the current uncontrolled availability?'. Does this imply a more complex trial design?

This research question is quite different from that which is currently conceived to underlie a randomised controlled trial. The most important difference is the importance placed on the context in which the controlled availability of heroin occurs, that is, 'What are the conditions under which...'

Stating the research question this way also begs questions about what we mean by 'harm'. We note that there are many different ways of constructing drug–related harm, and that drug related harm cannot be seen in isolation from drug policy. Some constructions reflect the chemistry of the drugs themselves. Others reflect characteristics of users. Still others reflect the illegal status of some drugs and society's responses to their illegality. The trial design will need to enable assessment of

harm reduction in a number of areas concurrently. However, the ability to change drug policy (or not) may be a major constraint on the extent to which drug–related harm can be ameliorated.

2   Issues for all RCTs that are especially relevant here

A number of issues applying to all RCTs form a backdrop to the consideration of service provision in a 'heroin trial'. They include the following:

- the RCT is most effective when it is testing a clear, preferably single, hypothesis. It is not so well suited to assessing the outcomes of complex interventions, including policy–related interventions like the controlled availability of heroin, particularly if the service delivery context in which the controlled availability occurs is different from that of conventional services;
- the design must be feasible. Heroin users are telling us what they want in a trial's service provision and the characteristics that will deter them from participation. Will they accept randomisation to an arm that does not provide heroin?;
- ethically, we cannot proceed with a RCT unless we are convinced that the design has sufficient statistical power to answer the research questions. How far do service provision considerations impact on this? Can we identify a design that can encompass the service provision realities?;
- the ethics of randomisation. A conundrum for all RCTs is that the whole basis for the RCT is that there is some evidence that the new 'thing' being tested has some chance of success at least for some people. Therefore, RCTs that are not conducted 'blind' can present difficult problems for service providers and participants who may have particular beliefs that one arm is likely to be of more benefit than another. They may feel that random allocation is unethical. An example may be useful here. What if an individual who is known to have done badly on methadone in the past is randomly allocated to the methadone group, when both that person and the service providers believe that the individual would do well on heroin? Almost certainly, allocation in the proposed trial will be unblinded and staff will be aware of participants' history, allocation and progress throughout the trial;
- generalisability. To obtain successful control, trials sometimes need to be far removed from everyday conditions, resulting in findings that are not generalisable. However, this research needs to be policy– and program–relevant and to answer important topical questions.

3   Principles of practice and how they contribute to health development

Without being prescriptive or comprehensive, we wish to give readers, particularly those with limited previous contact with heroin users and drug treatment agencies, some feel for the types of principles that we see as underlying a service provision model. Some are innovative, while others are seen in the operation of existing agencies. Some are well understood and apparently accepted by service staff and/or clients, but, for a variety of reasons, are not implemented. We then tie each group of principles to a concept of health development. The grouping is somewhat arbitrary as most service provision principles cut across each of the health development principles. The health development concepts arise from a broader program of research being undertaken by one of us (DL).

The first group of principles of service provision is that:

- clients have various ways of understanding their own drug use; these need to be respected and a service should not impose any particular theoretical understanding on the client;
- clients should determine their own desirable service outcomes;
- the quality of the relationship between staff and clients counts;
- the service will not judge the morality of the pleasurable aspects of drug use;
- individual differences will be recognised and fostered;
- nothing about the service should exacerbate the social marginality of many heroin users;
- clients have a right to have children.

These illustrate the health development principle that the service delivery is undertaken in a respectful collaboration with the people whose health it is (at both the clinical and community levels).

The second group of principles of service provision is that:

- the service should contribute to redressing the structural causes of the problems with which the clients present;
- there should be respect for the client's view of the link between the presenting problems and the structural causes of those problems;
- the concept of illegality based harm (as well as drug based harm) is valid;
- a plurality of reasons for drug use exist.

These illustrate the health development principle that the service delivery addresses the micro and immediate needs of sick care and public health in ways that also contribute to redressing problems understood in terms of macro level structures and longer time horizons.

The third group of principles of service provision is that:

- the service should foster social integration as it is perceived by the client;
- the service should encourage personal responsibility.

These illustrate the health development principle that the service delivery addresses the technical tasks of sick care and public health in ways that also strengthen our ability to manage existential challenges.

The fourth group of principles of service provision is that:

- society is providing a service to users and users have a responsibility to society in return;
- clients should be encouraged and given opportunities to be involved in the development of and management of the service, collaboratively with clinical and administrative staff;
- the principles of service operation should be explicit.

These illustrate the health development principle that the service delivery conceives the functions of sick care and public health as continuous with the broader domains of living a fulfilling life and contributing to building a better society.

We are seeking to use both the service provision and health development principles to develop a model of service provision for a trial that will contribute to harm reduction through impacting on a range of determinants of harm, rather than just providing heroin in a safe controlled manner. Obviously, a RCT could go ahead in a

manner that does not reflect these principles, but reproduces the stigmatising and controlling nature of too many contemporary services for heroin users. However, such an approach may well not contribute greatly to reducing harm.

4   Conclusion

Two possible outcomes from these considerations are to replace the RCT design or to further modify it.

With regard to modification of an RCT, we see a number of issues as problematic. Two issues that are particularly important, and are good illustrations of the difficulties inherent in an RCT, are client choice and individualised treatment.

Client choice. Some participants will want to move between the arms of a trial as it and they progress, and this needs to be allowed for. An example might be a person in the methadone only arm who has been doing well until faced by a crisis (such as the death of a loved one). The participant wants to start using heroin again to better numb the pain. Can they move arms or should they be forced back to street heroin?

Individualised treatment. Problems can arise when trial restrictions, or rules made to protect staff or clients, or laws conflict with what clients see as serving their own needs. To some extent these can be overcome with individualised treatment. Two examples illustrate this. First is the participant in the methadone arm who is continuing to use large amounts of street heroin. Should staff simply view this as a trial outcome or try to intervene? The second example is that in current methadone programs, a large percentage of clients have access to take–away doses. Take–aways are likely to be a problem for a heroin trial, both because of the perception of the risks of diversion and because trial participants may become the targets of violence. Should this mean that, for the sake of comparability, participants in the 'methadone arm' should also be refused take–aways? Should it mean that participants in the 'heroin arm' should have restrictions placed on their ability to leave the ACT for holidays, or to visit sick relatives, or even to stay at home in bed when they are ill?

We have raised for discussion issues concerned with the service provision features of a trial. This reflects our view that a trial is important in seeking to answer questions about the impact of the controlled availability of heroin and that the quality of service provision within the trial is also crucial for its effectiveness and for ethical and practical reasons.

The final proposal for the design and evaluation of a 'heroin trial' must address the dilemmas we have raised.

FEASIBILITY RESEARCH INTO THE CONTROLLED AVAILABILITY OF OPIOIDS

The Feasibility Research into the Controlled Availability of Opioids arose from a request to the National Centre for Epidemiology and Population Health (NCEPH) from the Select Committee on HIV, Illegal Drugs and Prostitution established by the Australian Capital Territory (ACT) Legislative Assembly.

A first stage of research, conducted in collaboration with the Australian Institute of Criminology (AIC), found that a trial to provide opioids, including heroin, to dependent users was feasible in principle. It was recommended that a second stage of feasibility investigations to examine logistic issues be conducted.

The first stage investigations examined illegal drug use in the ACT, the arguments for and against the controlled availability of opioids as reviewed in the literature, the current Australian political context for a trial, the role of interest groups in social controversies, legal issues, possible options for a trial, ethical issues, attitudes to a trial in the general community and among key interest groups (police, service providers, and illegal drug users and ex–users), and evaluation by a randomised controlled trial.

In addition, a proposal for a trial was developed as the starting point for the Stage 2 investigations.

The research which needs to be conducted to determine Stage 2 logistic feasibility can be divided into five areas:

- core information (for example, estimating numbers of users, determining relevant characteristics of ACT–based users, documenting the known information about the psychopharmacological and toxicological effects of opioids);
- information relevant to trial design and evaluation;
- information relevant to service provision;
- information about relevant legal, law enforcement and criminological matters; and
- community and key stakeholder acceptability of a specific trial proposal.

The Stage 2 research is also governed by the following principles:

- the research should have intrinsic value so that, regardless of whether or not a trial goes ahead, the research should be of value to treatment services or to drug policy generally;
- research should be conducted in all relevant disciplines and the disciplinary findings should be integrated to address the central problem;
- the process should involve to the greatest extent possible the key interest groups– illicit drug users, ex–users, service providers, police, policy makers and the community.

Stage 2 of the feasibility research into the controlled availability of opioids has many components. As significant advances are made in each particular substudy, we publish the results as a working paper, so that the information is available for discussion in the public arena.

PUBLICATIONS
Reports
* National Centre for Epidemiology and Population Health (1991), *Feasibility Research into the Controlled Availability of Opioids. Volumes 1 and 2.* NCEPH, The Australian National University, Canberra.
* Bammer, G. and Gerrard, G. (eds) (1992), *Heroin Treatment – New Alternatives*, Proceedings of a one–day seminar, Becker House, Canberra, November 1991.

Working papers
* Larson, A (1992), *Estimating the numbers of heroin users in the ACT,* Feasibility Research into the Controlled Availability of Opioids Stage 2, Working Paper Number 1.
* Bammer, G. (ed) (1993), *Australian drug markets research: What are we doing? Where are we going? What are the gaps? Proceedings of a one–day workshop held at the Scarth Room, University House, The Australian National University, Monday 22 February 1993*, Feasibility Research into the Controlled Availability of Opioids Stage 2, Working Paper Number 2.
* Sibthorpe, B., Sengoz, A. and G. Bammer in collaboration with The Youth Affairs Network of the ACT (1993), *Drug use and HIV risk among homeless and potentially homeless youth in the Australian Capital Territory*, Feasibility Research into the Controlled Availability of Opioids Stage 2, Working Paper Number 3.
* Bammer, G. and S. Weekes (1993), *Becoming an ex–user: Would the controlled availability of heroin make a difference?* Feasibility Research into the Controlled Availability of Opioids Stage 2, Working Paper Number 4.
* Bammer, G. (ed) (1993), *Does childhood sexual abuse contribute to alcohol, heroin and/or other drug problems? Proceedings of a one–day workshop held at the National Centre for Epidemiology and Population Health, The Australian National University, Tuesday 22 June 1993*, Feasibility Research into the Controlled Availability of Opioids Stage 2, Working Paper Number 5.
* Humes, G; Moloney, M; Baas Becking, F. and Bammer, G. in collaboration with Winnungah Nimmityjah (1993), *"It will kill us faster than the White invasion": Views on alcohol and other drug problems and HIV/AIDS risk in the Canberra/Queanbeyan Aboriginal community and on the suitability of a 'heroin trial' for Aboriginal heroin users*, Feasibility Research into the Controlled Availability of Opioids Stage 2, Working Paper Number 6.
* Butler, J.R.G. and Neil, A.L. (1994), *Econimic issues in a trial of the controlled provision of heroin,* Feasibility Research into the Controlled Availability of Opioids Stage 2, Working Paper Number 7.
* National Centre for Epidemiology and Population Health and Australian Institute of Criminology (1994) *Issues for designing and evaluating a 'heroin trial'. Three discussion papers,* Feasibility Research into the Controlled Availability of Opioids Stage 2, Working Paper Number 8.

Published papers
Hartland, N; McDonald, D; Dance, P. and Bammer, G. (1992), 'Australian reports into drug use and the possibility of heroin maintenance', *Drug and Alcohol Review*, 11, pp.175–182.
Bammer, G. (1992) 'A trial of controlled availability of heroin for the ACT?', in White, J. (ed) *Drug problems in our society: dimensions and perspectives* (Selected papers from the Window of Opportunity, First National Congress. An Intersectoral Approach to Drug Related Problems in Our Society, December 2–5, Adelaide), pp.57–62.
Bammer, G. (1993), 'Should the controlled provision of heroin be a treatment option? Australian feasibility considerations', *Addiction*, 88, pp.467–475.

McDonald, D; Stevens, A; Dance, P. and Bammer, G. (1993), 'Illicit drug use in the Australian Capital Territory: Implications for the feasibility of a heroin trial', *Australian and New Zealand Journal of Criminology,* 26, pp.127–145.

\#   Bammer, G; Douglas, B; Moore, M. and Chappell, D. (1993) 'A heroin trial for the Australian Capital Territory? An overview of feasibility research', in Heather, N; Wodak, A; Nadelman, E.A. and O'Hare, P. (eds) *Psychoactive drugs and harm reduction: from faith to science* (Selected papers from the Third International Conference on the Reduction of Drug–related Harm), Whurr, London, pp.137–150.

\#   Ostini, R; Bammer, G; Dance, P; and Goodin, R. (1993) 'The ethics of experimental heroin maintenance', *Journal of Medical Ethics*, 19, pp. 175-182.

Newsletters

\#   Newsletters reporting project results are also published from time to time.

\# These publications are available free from:	* These publications are for sale through:
Dr Gabriele Bammer	**Bibliotech**
Feasibility Study Co–ordinator	The Australian National University
National Centre for Epidemiology &	ACT 0200
Population Health	
The Australian National University	
ACT 0200	
Phone: (06) 2490716	
Fax: (06) 2490740	