Home Office

BUILDING A SAFE, JUST AND TOLERANT SOCIETY

The feasibility of conducting an RCT at HMP Grendon

Siobhan Campbell

Home Office Online Report 03/03

The views expressed in this report are those of the authors, not necessarily those of the Home Office (nor do they reflect Government policy).

Contents

Executive summary **4**
1. Introduction **7**
Background 8
The feasibility study 9
2. HMP Grendon **11**
The regime 11
3. The feasibility of an RCT **15**
Randomisation 15
Statistical Power Analysis 16
Numbers 17
Case-flow 17
Volunteers 20
The control group 21
Timeframes 22
Summary 23
4. Conclusions **25**
5. References **26**
Appendix A: Questions and answers **28**

Executive summary

Background

The regime at Grendon is of interest to the Dangerous and Severe Personality Disorder (DSPD) research and development programme for a number of reasons. First, there is likely to be some degree of overlap between Grendon's population and the DSPD population: it is already known that 47 per cent of Grendon inmates score 25 or above on the PCL-R (Psychopathy Checklist). Second, those with high PCL-R scores tend not to be treated by the Prison Service's accredited programmes, leading to a paucity of evidence about high PCL-R scorers response to treatment. Third, there is some ambiguous evidence as to the efficacy of Grendon when considering long-term outcomes: some studies have shown it can reduce reconviction, some that it has no impact and others that it may in fact relate to poorer long-term outcomes. However, none of these studies can be taken as conclusive evidence because of uncertainty surrounding the comparability of the control group. For all these reasons, a robust, reliable and unambiguous study is needed of Grendon.

A Randomised Controlled Trial (RCT) is proposed as the best way of evaluating Grendon because it can deliver the required robust, reliable and unambiguous results. Through randomisation, the equivalence of the treatment and the control group can be guaranteed. In addition, any change at the end of the experimental period can confidently be attributed to the treatment at Grendon, and not to some third factor such as maturation or internment.

HMP Grendon

There are a total of 230 beds at Grendon, with a maximum of 200 inmates undergoing treatment at any one time. There are five treatment wings, each with 40 inmates, which are then divided into smaller therapy groups of between five and eight. Approximately 115 prisoners enter Grendon each year.

To enter Grendon a prisoner must be over 21; not on psychotropic medication; drug-free; have at least two years still to serve; have no outstanding appeal or parole proceedings; and be willing to participate in therapy. Historically Grendon was for personality disordered offenders, and individuals needed a medical referral to apply. However, over the last 18 months the application process has been opened up considerably and prisoners can now apply directly to Grendon. In addition, all mention of personality disorder has been omitted from the newly produced leaflets and application forms.

On entering Grendon, prisoners undergo an eight to 16 week assessment. At present this is not formalised, and few standardised assessments are conducted. However, with the appointment of a new therapist to lead the assessment wing it is hoped more rigorous clinical and psychometric assessment procedures will be brought in. After this time, the prisoner is assessed as appropriate for a certain wing and a certain therapy group. Once chosen and placed, the prisoner will not move from that wing. The average length of stay is between one and two years, although some stay longer. Overall, only five per cent of the population stay over three years.

The therapy is continuous, with three small group sessions a week and two large group sessions. Each wings is under the supervision of a different therapist, all of whom have quite different beliefs and work semi-autonomously. For this reason each wing tends to be run quite differently.

A number of tests are currently run on Grendon's population by the on-site research team. However, those planning an RCT would have to think carefully as to what tests are necessary because the same tests will also need to be conducted on the control group who will be dispersed across the prison system. At present, therapists only judge 44 per cent of those who start therapy to have either partially or fully meet their therapy aims, although this is a

relatively arbitrary measure, perhaps reflecting more about the perceptions of the therapists than the success of treatment.

The feasibility of an RCT

A number of difficulties were identified as possible hurdles that would need to be overcome if an RCT was to be possible. One of the most pivotal is support for an RCT from within Grendon and also from within the Prison Service. Many people within Grendon recognised the potential benefits of an RCT in providing unambiguous results; raising the profile and status of Grendon; and providing support for the accreditation process.

There were some doubts about an RCT. However, by minimising the impact of the trial on the everyday work at Grendon, and by not interfering with professional decision-making or clinical judgement, it is hoped that many of these obstacles can be overcome.

Another potential obstacle is that there is currently only 25 people on the waiting list to enter Grendon. Although this is suitable for their day-to-day work, it is far too small a group from which to randomise. These men will enter Grendon as beds become available in order to maintain working to capacity. The only way to ensure randomisation is for demand to greatly exceed supply. Previously, roadshows have managed to increase the demand for places at Grendon: this technique could be used again. In addition, if there was wider Prison Service support, other prisons could be encouraged to inform prisoners about Grendon and support them in their applications.

An analysis of the numbers required in the study revealed that the minimum number possible is 64 in each group to detect an effect size of 0.25 with a statistical power of 0.80. However, this is very much the minimum required. More practically, it is possible to assess the statistical power of the study given that the average number of 115 participants passes through in a year. If there was again an effect size of 0.25, the statistical power of the study would be a very acceptable 0.97. To calculate a working figure a power of 0.80 and an effect size of 0.20 was estimated resulting in a desired sample size of 100 in each group.

However, when the case flow through the trial is considered, potential difficulties with these numbers are revealed. These centre round the fact that there tends to be a very high dropout rate from Grendon: as discussed earlier, only 44 per cent of those who enter Grendon either partly or fully meet their therapy aims. Using the case flow estimates only 33 per cent of those initially assessed as suitable according to their application and therefore randomised were expected to successfully complete their treatment at Grendon. It is, of course, desirable to increase this proportion; however, even if this was possible it has to be kept in mind that there will be a large proportion who will not complete the treatment.

A number of scenarios were considered using the case flow estimates of the ‘worst case’ scenario for dropouts in each case. The first used a sample of 100 and resulted in only 33 successfully completing the treatment. The second scenario ensured 100 successfully completed, but needed 673 to have initially volunteered to enter the study. The compromise scenario plotted 200 volunteers in each of the groups, and 66 successfully completing. These scenarios clearly illustrate the need to carefully monitor the dropout rates throughout any evaluation of Grendon.

The final obstacles highlighted concern the control group. There is a need to take great care when considering the control group. To volunteer for a treatment and be refused on the basis of chance may be quite damaging for those in the control group, and there is a possibility they decompensate and engage in more problematic behaviour than they would otherwise. Randomising early might help minimise such feelings. In addition, there may be ethical problems surrounding permanently excluding people from Grendon. To overcome these problems offering another ‘treatment’ or form of support would be advisable. In this way prisoners are being randomised between two potentially beneficial treatments rather than between treatment and no treatment. The report suggests a personal advisor as a potential control condition.

Conclusions

The report concludes that it is indeed possible and desirable to conduct an RCT in Grendon. However, it is acknowledged that there are a number of obstacles that would first need to be overcome. The discussions of such a trial must now be passed onto the people who can implement such changes. Changes to the senior management team at Grendon have resulted in a new team who should be approached with the results of this study.

1. Introduction

One of the key aims of the DSPD research programme is to investigate potentially successful interventions with those who might satisfy the criteria of DSPD. Part of this will be to evaluate the interventions to be piloted in HMP Whitemoor and Frankland and Broadmoor and Rampton High Secure Hospitals when they are rolled out and to evaluate the efficacy of the newly developed psychopathy programme.[1] However, in addition to these studies, and in order to develop a 'What Works?' evidence base, it is also important to ensure that all potentially successful interventions with dangerous individuals and those with severe personality disorder are fully explored with the highest quality research possible. The work conducted in HMP Grendon is certainly of great interest to the DSPD research and development programme.

The model of intervention currently being used at HMP Grendon is of interest for a number of different reasons, but most importantly because of the ambiguity of previous research results on its outcome and the nature of its population.

A series of evaluations and studies has been conducted on the efficacy of the regime since it opened in 1962. Indeed a book has been compiled detailing much of this research (Shine, 2000). However, the nature of these studies, while being encouraging on the whole, have not given conclusive answers as to whether the regime at HMP Grendon is effective or not in reducing long-term reoffending. In addition, the currently available evidence is rather dated and may not, in fact, reflect the current regime or population.

Another important aspect of the regime at HMP Grendon relates to the nature of its population: at least one inmate who has been through the Whitemoor DSPD has since volunteered for Grendon and a high proportion of psychopathic offenders have been treated over the years. Recent research has indicated that the proportion of psychopathic offenders is very large (Gray *et al*, unpublished report to the Home Office, 2002), mirroring earlier findings by Hobson and Shine (1998). Hobson and Shine concluded that in 1995 27 per cent of Grendon's population scored 30 or over on the Psychopathy Checklist (PCL-R, Hare, 1991) indicating a severe form of psychopathy. Gray *et al* reached a similar figure, although discussed the lower, British cut-off of 25. They indicated that almost half (47%) of the current (2002) population can be assessed as psychopathic by scoring 25 or more on the PCL-R.

There is therefore a clear need to look at the way psychopaths are responding to HMP Grendon's regime, and whether their outcomes parallel those of inmates not assessed as psychopathic. It is also important to know whether psychopaths in Grendon pose a greater or lesser risk than psychopaths contained within the rest of the prison estate.

There are a number of ways the regime at Grendon can be evaluated. Over the years studies have been conducted using before and after designs, or comparisons have been made with those treated at Grendon and those on the Grendon waiting list. However, as enlightening as many of these studies have been, their results have been mixed and it is possible to argue that the results or null-results obtained are due to some other factors other than the experience at Grendon. For example there might be a bias towards those who enter Grendon compared with those left on the waiting list: Grendon people may be more or less likely to reoffend, or more or less likely to respond to treatment. Those on the waiting list may be more indecisive, or more mobile and move around the prison system. They may be more likely to become involved in adjudications or other disciplinary matters or differ in terms of their psychopathology. For these reasons, comparisons between the two groups may be unjustified. In addition, before and after results could relate to many other factors, for example maturation or length of sentence, and should not be read in isolation from a truly comparable control group.

[1] The psychopathy programme is currently being developed by the Offender Behaviour Programme Unit in the Prison Service. It is designed specifically to treat psychopathic offenders who may not be treated in other parts of the prison service with the aim of lessening their risk of seriously reoffending. It is due to begin piloting in April 2004.

For this reason it is suggested that an RCT is probably one of the most reliable ways to evaluate HMP Grendon and produce unambiguous conclusions about the causal relationship between participating in HMP Grendon and longer-term outcomes. Put simply, an RCT will ensure the comparability of the treatment and control group, allowing any differences observed between the groups to be confidently attributed to the treatment individuals are receiving at Grendon.

It could be argued that equally robust results could be obtained through other forms of evaluation. However, ensuring the comparability of the treatment and the control group would be a difficult and more laborious process than randomisation, without the guarantee of comparability at the end of the process. Despite this, many of the issues discussed in this report will also be applicable to other types of evaluation.

Background

This report follows on directly from Farrington and Jolliffe (2002). It is intended that they are read together and complement each other. For this reason, much of the historical background of RCTs in Britain will not be discussed in this paper. However, in summary, their study concluded that there is a paucity of RCTs in the social sciences in the UK, with only five being attempted in prison settings over the last 40 years.

Their paper concluded that an RCT is indeed possible and desirable in UK prisons (specifically HMP Whitemoor), if a number of criteria could be satisfied. They pointed to the issues surrounding:

- the number of participants. If the number in the experimental group is externally limited, the number in the control group needs to increase dramatically to compensate (they estimated at least 300 controls to compare with a 50 person sample)
- the length of the treatment. The longer the treatment, the fewer people being treated over the course of the study
- assessing all the participants, including the control group. This is arduous and difficult, especially for the control group which may be spread out across a number of prisons.

Other issues, which may have been expected to prove difficult, proved less so. In discussion with both an ethics committee and the Parole Board Farrington and Jolliffe found that it was likely that an RCT would not prove problematic to an ethics committee, and it was unlikely to give those in the treatment group an unfair advantage with the Parole Board.

However, there are also some important factors specific to HMP Grendon that should not be overlooked. First, for the last 40 years HMP Grendon has been run along the lines of a democratic Therapeutic Community (TC). There is a great deal of evidence that TCs such as that run in HMP Grendon have a positive impact on a range of psychiatric, personality and psychological variables (Gunn *et al*, 1978; Genders and Player, 1995). However, there are also studies which have indicated that some forms of TC can in fact make psychopathic offenders worse (Rice *et al*, 1992). Rice *et al* found that while the TC regime they examined was effective for non-psychopathic offenders, psychopaths were significantly more likely to re-offend than controls. This is not conclusive evidence that TCs do not work for psychopathic offenders: the TC examined in this study had some very unusual features, and cannot be taken as representative of TCs in general. However, it does point to the care that needs to be taken in the treatment of psychopaths and the need for rigorous evaluation of all work with this group.

At present, there is no evidence that psychopathic offenders have any different long-term outcomes to non-psychopathic offenders at Grendon, although previous research at Grendon has found that high PCL-R scores are predictive of highly disruptive behaviour during treatment sessions and on the wings (Hobson *et al*, 2000). This again suggests that a thorough evaluation of psychopathy and outcomes at HMP Grendon is needed.

There is also a debate as to the type of TCs that would be suitable for psychopathic offenders; Losel (1998) suggested that the ideal TC should be highly structured with clear rules, regulations and duties to help staff retain control and reduce the opportunity for manipulation. He suggested that low structured forms of TC are less successful and can be counterproductive. It could be argued that the fact that Grendon is run along democratic grounds might be allowing the psychopathic members of the community an opportunity to undermine the regime. However, equally it could be argued that the regime at Grendon is also very structured. Again an RCT would help unpick the effect of this specific regime on this population.

Finally, Grendon's population is not a simple division between psychopaths and non-psychopaths; there are many other aspects of the population that would be interesting to explore. For example, TCs have been shown to be effective in the treatment of some forms of personality disorder. Chiesa and Fonagy (2000) found that a hospital based TC was effective in the treatment of borderline personality disorder. It would be important to know whether Grendon is more effective in treating some forms of PD above others.

On the whole, previous research into the work of Grendon has either been positive or indicated that there has been no effect either one way or another. However, issues surrounding the control group used in these studies and whether it is in fact directly comparable to the treatment population, are still a subject for debate. Taken together, all the previous evidence points to no firm conclusions. The results are ambiguous and there is a strong need for clear, unbiased results that will be accepted by the whole of the academic community.

The feasibility study

HMP Grendon has an on-site research team which is constantly monitoring and evaluating progress. Over the years various groups have written no fewer than 62 articles about the regime at Grendon and the outcomes of the regime. Grendon seems, on first inspection, not to need yet more evaluation. The question therefore needs to be asked: why do we need an RCT at Grendon?

There are many reasons why such an evaluation would be valuable:

- As discussed in the introduction, the evidence in the numerous published and unpublished research articles is inconclusive. Some indicate improvement for those leaving Grendon against a control group (Marshall, 1997; Taylor, 2000), others have reported no difference (Robertson and Gunn, 1987). There is also some emerging evidence that suggests the likelihood of reoffending on leaving prison might actually increase if the prisoner has attended Grendon (Newton, unpublished). Conclusive causal evidence is therefore needed to resolve the ambiguities in this current evidence base.
- The comparison group that has typically been used to compare the HMP Grendon population is usually made up of those who applied to enter Grendon, but, although eligible, did not. This waiting list comparison group is typically small (much smaller than the population it is being compared with). Its direct comparability with the Grendon population can also be doubted in that it is not firmly established why these individuals did not enter Grendon. Some have supposed that these people did not enter because no place became available soon enough or because the individual was given parole earlier than expected (Marshall, 1997). However, it may also be the case that they did not enter Grendon because they had drug problems; became involved in disciplinary offences; or became involved in appeal or parole proceedings, making this group substantially different from the Grendon population. Comparing the effect of Grendon with a truly comparable group on every factor (both measured and unmeasured) is one of the few ways these problems can be overcome.
- In many prison-based programmes, psychopathic offenders are excluded because of the fear that the programme may make them worse. However, in Grendon, as discussed earlier, roughly half of their population has a PCL-R score of 25 or more. For this reason,

it is essential we obtain as much good quality information as possible on the outcome of participating in Grendon on psychopathic offenders. Such information could in turn help advise the admissions policy for Grendon, and also give valuable information to other prison-based initiatives such as the psychopathy programme which is currently being developed.

- HMP Grendon is currently going through the process of accreditation. Part of the accreditation process involves the production of robust evidence in support of the programme. An RCT is the 'gold standard' of evaluation, and will be able to give the most reliable and robust information to this process.
- Finally, the admissions process has opened up considerably over the last few years. For this reason, it is reasonably to assume that the population profile has also changed over the same period. It is feasible that the prisoners currently in HMP Grendon might differ considerably from earlier cohorts and be less or more amenable to therapy.

2. HMP Grendon

HMP Grendon for a long time was a unique provision for personality disordered prisoners within the prison estate. It is a prison, run along the lines of a Therapeutic Community (TC) on completely democratic lines. In addition, it also houses its own research unit.

More recently, other prisons or wings run along therapeutic lines have been established within the prison estate, but Grendon's place has remained unique through its long history and strong advocates. These other establishments will also be evaluated in time, but are not directly comparable with Grendon because of different entry criteria. For example, the largest, HMP Dovegate, refuses entry to anyone with a PCL-R score over 25.

The regime

Entry criteria

HMP Grendon has a total of 230 beds, with five residential wings with a maximum of 40 people per wing and an assessment unit holding a maximum of 30, although in practice there are 25 beds. Within the wings of 40 people, the prisoners are split into smaller groups of between five and eight. On a day-to-day basis, the number on each wing is typically between 34 and 36.

The criteria for entering HMP Grendon have changed considerably over recent years. Previously, application was dependent on a medical referral (a 1080). This limited applications, and resulted in a few establishments sending a disproportionate number of prisoners because their medical staff were keen to refer.

There are now new criteria for entering Grendon which are more inclusive. Prisoners are sent information directly throughout the prison estate and the information and application pack has been upgraded and is now presented in booklet form. If prisoners are interested in applying to Grendon, they can now nominate themselves, although they still need the support of staff members.

The only specific criteria remaining for entry to Grendon are for the prisoner to:

- be over 21
- not be on psychotropic medication
- be drug free
- have at least two years still to serve
- have no ongoing appeals or parole procedures
- be willing to participate in therapy.

Previously, prisoners referred by their medical officer had to be assessed as having a personality disorder. However, now it is feasible for non-personality disordered prisoners to refer themselves. Grendon is a Cat B prison but it is not unknown for prisoners to apply or be referred from Cat C or even D prisons, indicating their commitment to therapy.

Once the application form is received by Grendon the head therapist assesses whether the prisoners should come to Grendon for assessment or not. However, in practice, few are refused.

The emergence of the other TCs prisons in the prison estate, most notably HMP Dovegate, has led to a slight fall in the number of applications made by prisoners to enter Grendon. In addition, because Grendon has yet to be accredited, prisoners might feel that it would be

more beneficial for their progress through the prison system to remain in prisons that can offer the Prison Service's accredited programmes rather than to apply to Grendon. The change to the entry criteria and the improved documentation were designed to combat this falling waiting list. In addition to this, in 2001 a series of roadshows was conducted to encourage applications to Grendon. These roadshows were led by multidisciplinary teams (probation officers, therapists, uniformed officers, prisoners and the population manager) and were held wherever their host prison had room. Staff in Grendon did feel that the roadshows were successful in increasing the number of applications to Grendon, but they are suspended at present because of staff shortages.

There is now a contract between the sending prison and HMP Grendon that they must take prisoners back at any time. Before this arrangement, the authorities at Grendon had problems placing disruptive prisoners, with the sending prison refusing to take them back. However, it is hoped those successfully completing their therapy at HMP Grendon will be able to move to a lower category prison on leaving.

Admissions

Currently there is a throughput of roughly 115 people a year. Prisoners are assessed on F wing. There are 25 beds, and a maximum of three admissions a week. Assessment takes between eight and 16 weeks, depending on the individual.

The methods of assessing new inmates use representatives and therapists from all the communities. Prisoners are assessed as to whether they are amenable to therapy and whether they would fit into specific groups. This tends to depend on the type of offender as well as other factors, for example, it is important to have a mixed group: a group of armed robbers will simply feed off each other. In putting together the small groups, it is hoped that a mix of people will learn and challenge each other.

People will fail the assessment primarily through their behaviour: by committing sexual or violent acts or drug or alcohol infringements. However, others will simply be rated as 'unsuitable for group therapy'. In practice, the majority leaving at this stage will be because of the infringement of rules. However, this does not prohibit them from reapplying. It is understood that they may simply have come at the wrong time in their lives and so most who reapply are reaccepted.

Currently assessment is not standardised with different therapists having different perspectives. Because of this, any evaluation cannot guarantee what information will be routinely available and little of the time in the assessment unit is spent engaged in formal assessment. Most of the formal assessment to date has been conducted by the research team rather than relying on what may be available at the end of the assessment process. However, this is due to change with the appointment of a therapist to lead the assessment wing who intends to bring in more rigorous clinical and psychometric assessment procedures.

The wing

Once chosen by a therapist for a specific wing, prisoners visit the wing and meet the wing chairman. Two to three days afterwards they join the group, although they do not start therapy immediately. Once on the wing and in a particular group, the prisoner does not move. The average length of stay is between one and two years, although some stay more than four or five years.

Wings A, B, C, and D are similar. Wing G is specifically for sex offenders. This is not to say that there are no other sex offenders at Grendon: there are sex offenders in all four of the other wings. However, G wing is reserved for the most serious offenders who may need more specific work.

There is no rule 45 at Grendon (segregation), all prisoners mix throughout the prison. In addition, there is no work as in other prisons. This is because the time spent in therapy does not leave enough time for traditional work.

Therapy

Everything that happens in Grendon happens democratically, even to the degree that representatives from each wing converge on the governor and give him regular feedback, both positive and negative.

There are therapy sessions every morning for all prisoners. On Monday, Wednesday and Friday, prisoners meet in their small group of between five and eight. On Tuesdays and Thursdays the entire wing community meets. This occurs every week, with the exception of three or four 'therapy holidays' a year where therapy stops for two weeks. This can be a chaotic time, with the prisoners losing their normal structure. However, it allows them practice for everyday living without daily therapy.

Each treatment wing has a therapist who is in charge of treatment on the wing and a senior psychologist. Each community works autonomously. Because of this the five communities are very different and there tend to be some divergent opinions between therapists, resulting in no coherent single belief shared by the psychotherapists at Grendon. One therapist reported that the only consistency between the therapists was 'getting things out in the open'.

In each community, prisoners are taught to talk and trust and 'break down the barriers' and it engages them in an increasingly social process. Some therapists do not believe in the more standard cognitive behavioural therapy (CBT) approaches used in other parts of the Prison Service. One therapist stated that the intellectualism of a problem has always been seen as a defence against emotion in therapy, but he believes this is what CBTs are doing. He felt that emotions, not cognitions, drive actions and therefore this is what should be treated.

In therapy, all risk factors are challenged, especially with lifers. Often prisoners have previously completed Prison Service Offender Behaviour Programmes (OBPs) but if there are still areas of need, these areas are gone over again. One therapist stated that Grendon works well with psychopathic offenders because 'they have had a cold and callous upbringing – they need very strong support networks' which is what he feels Grendon gives. He felt there was a perception that psychopaths charm their way through the system, but he felt that this theory did not hold up in Grendon because they were encouraged to expose and confront their belief systems.

Outcomes

There is a lot of evidence that Grendon does change people. A history of psychometric tests and other short-term proxy measures indicate positive change (eg. Newton, 1998). There are good outcomes of self-report questionnaires, adjudications etc, with prisoners who spend time at Grendon making improvements in many of these measures between entering and leaving the prison. However, the evidence that is called most into doubt is reconviction data. As stated earlier, there is no conclusive evidence that spending time in Grendon has a long-term impact on inmates' likelihood of reconviction.

The tests traditionally given to new inmates at Grendon by the research team are:

- Raven's standard progressive matrices
- Culture free self esteem inventory (Battle)
- Blame attribution inventory (Gudjonsson)
- Eysenck personality questionnaire revised
- Eysenck adult impulsiveness, venturesomeness and empathy scale

- Hostility and direction of hostility scale (Foulds and Caine)
- Grendon problem checklist.

These assessment measures are currently being revised and different measures and risk assessments are being considered. For example it is likely that in the future each inmate will be assessed using the PCL-R. However, in a full-scale RCT the outcome measures used will have to be considered carefully: only those regarded as essential should be used. This is because it will be very difficult to assess those in the control group across the prison system, and any tests conducted would need to take as little time as possible.

There is currently a culture of non-participation in re-testing. Because of this the research team at Grendon are asking new inmates to sign an agreement to be re-tested when they enter Grendon as there is no problem with prisoners agreeing to participate on first entering Grendon.

At present, when someone leaves Grendon, staff are asked to assess whether they think their therapy aims have been met: 56 per cent are reported not to have met their aims; 25 per cent have partially met them and 19 per cent have fully met them. However, most who stay 18 months or more meet their aims.

Interviews suggested that there are some in Grendon who do not like the measurement of inmates: they feel it can get in the way of working and can lead to a different focus to that of the therapy. However, this was a view not universally held. Some commented that there is no coherent view on what Grendon is trying to do and how to measure it. There are differences in opinion with some in Grendon focusing on the prisoners, criminogenic needs, and other focusing on more person-centred difficulties. This could potentially lead to difficulties in obtaining standardised measures of need and progress. In addition to this, it may be difficult in getting agreement as to what the tests that should be conducted are. For example, one interview revealed quite a lot of hostility to the PCL-R, although this may be a criticism towards the PCL-R as a clinical tool, rather than as a research tool. Criticism of the PCL-R because of its static nature may be overcome if the study is trying to demonstrate differences in reconviction rates between people with equally high scores based on the treatment they received.

One of the therapists highlighted the need for new measures. Grendon is not just targeting prisoners' criminogenic needs; it is also targeting other, more personal and social factors. For this reason factors such as quality of life should also be measured.

3. The feasibility of an RCT

In the abstract, an RCT is easy to design: it is simply a matter of calculating the numbers required and the method of randomisation. However, this is obviously not the case in practice. In order for an RCT at Grendon to be practicable, a number of problems must first be overcome and a number of conditions met.

Some of the difficulties envisaged in conducting an RCT include:[2]

- There would need to be support from a number of different individuals and agencies including the Prison Service and all interested parties in Grendon.
- Conducting an RCT would interfere with normal processes. It would necessitate redesigning allocation and selection.
- Fundamental to an RCT being feasible is to increase the demand for places – there are currently too few people on the waiting list.
- Because of the pressure on beds in the wider Prison Service, it is thought that even if there were sufficient numbers on the waiting list, they would be given beds in Grendon as soon as they became available.
- If some in the control group are, in fact, DSPD it is likely that they will come into contact and be eligible for DSPD services when they come on stream in the future. This would severely affect the study.
- At present 40 per cent of those on assessment and 50 per cent of those in the wings are lifers (they tend to stay longer in Grendon). If one of the most important outcome measures for Grendon is reconviction, then this may strongly affect outcome.[3]

The first and most fundamental of these obstacles is organisational support. If there is no support from managers, therapists and administrators at Grendon, then there is little chance of an RCT succeeding. Even if the project did get off the ground, there would be little likelihood of success in practice without this kind of support.

Interviews with some key players at Grendon revealed mixed support. The acting governor was supportive of the idea: an RCT would raise the profile of Grendon and deliver strong evidence on the success or otherwise of the treatment. It would also help support the current application for accreditation if it could be shown conclusively by way of an RCT that Grendon was effective. He was also fairly confident that this would be a view that the incoming governor would hold, and thought it likely he would also give general support to the idea, although there is no concrete evidence that this would be the case. The research team was also supportive of the proposal, understanding the benefit of the robust evidence an RCT would bring. However, they were aware of the practical hurdles that would need to be overcome. The therapists as a group were more wary of an RCT, although supportive of research in general.

Randomisation

As stated earlier, there was unease amongst some of the staff at Grendon about the idea of an RCT, with some contending that even medical trials were flawed. The scepticism towards random allocation is potentially problematic. It would be important to give the staff, particularly the clinicians, as much information as possible to ensure their concerns are acknowledged and addressed where possible.

[2] There are a number of question and answers on frequently asked questions about RCTs in appendix A.
[3] By definition lifers will be usually in prison longer than other inmates. Lifers released on life-licence will be under strict supervision and may be recalled (rather than reconvicted) at even minor infringements. This will affect these individuals' ability to reoffend and typically, lifers have very low reconviction rates (between 1972 and 1994 9% were reconvicted within two years). It will be important, therefore, for there to be a comparable number of lifers in both the treatment and the control groups and for them to be analysed separately.

Thus, concern about the RCT interfering with therapy can easily be dealt with by randomising early, before the participants enter Grendon, ensuring that the work of the therapists, psychologists and prison officers is disturbed as little as possible. This means that, although the assessment wing would seem the obvious place from which to randomise from an experimental design perspective (comparable assessments would be completed for all prisoners and only those deemed appropriate for treatment would be randomised) from a practical perspective, it would not be possible. The assessment unit is where individual therapists observe and interact with new prisoners and make decisions about their suitability to fit into a team/wing. Interfering with their professional judgement at that stage would be absurd. In addition, there is a strong risk that prisoners assessed and deemed suitable for treatment but refused admission on the basis of chance would (a) decompensate and become involved in more problematic behaviour than if they had never been involved in the trial; and (b) refuse to participate in the remainder of the trial.

For these reasons, randomisation as early as possible is recommended. The most practical stage would be once the prisoner had volunteered and been assessed as suitable for assessment. At this stage there could be an independent process that would assign prisoners to either Grendon or the control condition. Only those in the treatment group would therefore enter Grendon or be assessed for entry to the wings.

Statistical power analysis

Following on from the discussion of statistical power analysis in Farrington and Jolliffe (2002), we can use the same theory to specify what sample size is needed for given values of the effect size, statistical significance and statistical power. Using the same assumptions, statistical significance was set at p=0.05. The effect size, again using the same assumptions as Farrington and Jolliffe, based on the meta-analysis by Redondo *et al* (1999) of the influence of European treatment programmes on recidivism. This estimated an average effect size of 0.24. With a confidence interval of 95 per cent, the average expected effect size was estimated at between 0.20 and 0.29 over all subjects and all programmes. The 0.24 average corresponds to a change in the percentage reconviction rate from 50 per cent to 38 per cent (Lipsey & Wilson, 1998).

As power is maximised if the sample size in the treatment group is the same size as the control group, the following discussion will assume equal sample sizes. The estimated numbers required in the study using these assumptions – an effect size of 0.25 and a statistical significance of 0.05 – and given a sample size of 115 (the average number to pass through Grendon in a year) results in a power score of 0.97 (Table 3.1).[4]

Table 3.1: Power calculations for different effect sizes and a sample size of 115 in each condition.

n	Alpha	Effect size	Power
115	0.05	0.10	0.33
115	0.05	0.25	0.97
115	0.05	0.40	1.00

However, it is rare for any study, and especially an RCT, to go according to plan. The initial recruitment of participants and their retention in the trial are the most common problems encountered with this type of study. However, power calculations can also be made to understand what the smallest sample size possible is while maintaining all the previous assumptions. In this case 64 individuals would be required to detect a minimum effect size of 0.25 and a power of 0.80 (Table 3.2). If the effect size was smaller than this, far more participants would be required, for example to detect an effect of 0.1 at least 310 participants would be required in each group. The most realistic estimate is that at least 100 participants

[4] The calculations were conducted using the statistical package ‘Power and Precision’. Specifically, power calculations were conducted on one-way analyses of variance with two levels. Various effect sizes and power levels were used, but the alpha level remained 0.05.

would be required in each of the groups in the study to identify an effect size of 0.20 with a power of 0.80.

Table 3.2: The sample size required for to detect different effect sizes with different power cut-offs (alpha =0.05)

Power	0.10	0.20	0.25	0.30
0.70	310	79	51	36
0.80	394	100	64	45
0.90	527	133	86	60

Note: A power level of 0.70 means that the researcher would expect to falsely obtain a negative result in 30% of cases. Thus the higher the power, the less chance of making this false negative error.
An effect size of d=0.10 represents a 5% reduction in recidivism; d=0.20, a 10% reduction and d=0.30 a 15% reduction (Farrington and Jolliffe, 2002).

Numbers

The implications of the statistical power analysis are that at least 100 participants would be required in each of the groups. If more detailed analyses were to be conducted, for example looking at the progress of those with high PCL-R scores or the progress of lifers, even higher numbers would be required to ensure each group contained 100 individuals. Would it be possible to obtain this number of volunteers?

On average 115 prisoners enter HMP Grendon each year. In order for the results of the study to be reliable, each person entering Grendon during the course of the study should be a participant. If not, then the sample may be biased and therefore not representative of the population as a whole. In this case, even if a significant result is found, if only a very small proportion of the original sample volunteered, no generalisation could be made from this group to the population. Ensuring as many people volunteer for the study as possible is essential.

However, it is unethical to coerce prisoners to participate, and there will be natural dropout from the study. In addition, RCTs are notorious in that it is rare to obtain the numbers originally anticipated in the original timescale. For this reason if the study was designed to continue over two years, even if only 50 per cent of the original population remained in Grendon and in the study, more than the originally required 100 participants will still be obtained.

However, perhaps the most ethical way to ensure the numbers required is to recruit individuals until the target number is established. For example, it is probable that the numbers required would be achieved within two years, so it is perhaps best to state that the study will continue until the required number of individuals is achieved, or at two years, whichever is first.

Case flow

Figure 3.1 describes a possible flow of cases into the RCT. Worst-case scenario dropout rates have been estimated at each stage of the process to establish whether, even with the most pessimistic progress, an RCT would be feasible. All analyses of RCTs are on an ‘intention to treat’ basis. Thus, although there will be those who drop out, in order to ensure equivalence across the treatment and control groups, all those who were originally in the treatment group (N4) will be analysed as if they were treated. For this reason it is desirable to minimise dropouts as much as possible at stages N5 and N6.

The smallest sample scenario [5]

Working from the treatment group (N4), the statistical power analysis revealed that at least 100 would be required in this group for an effect size of 0.20 to be reliably detected. If this was the case, and the dropout rates were those estimated in Figure 3.1, then 222 volunteers would be required to apply for Grendon at N2: it is estimated that 22 (10%) of these individuals would not be eligible. One hundred individuals would be randomly assigned to the control group and 100 to the treatment group (N4) (Table 3.3).

Working forward, it is likely there will be a number of dropouts after randomisation. Of the 100 people entering assessment, it is estimated that 75 per cent will be assessed as suitable and progress to the treatment stage, leaving 75. Finally, it is likely that there will be a large number of dropouts from the treatment programme after assessment, mainly through the infringement of Grendon's rules. At present 56 per cent of inmates are deemed by the therapists not to have completed successfully (although this is obviously a very subjective measure of success). Thus, if only 44 per cent complete successfully, the final number completing (N6) will be just 33.

[5] The smallest possible figure estimated in the statistical power analysis was 64. However, this was considered too small considering the potential dropout rates (only 21 would successfully complete). The study therefore assumes 100 is the smallest desirable sample the RCT is likely to work with.

Figure 3.1: Participant flow chart for RCT (minimum numbers required)[6]

It is unknown how large N1 may be.

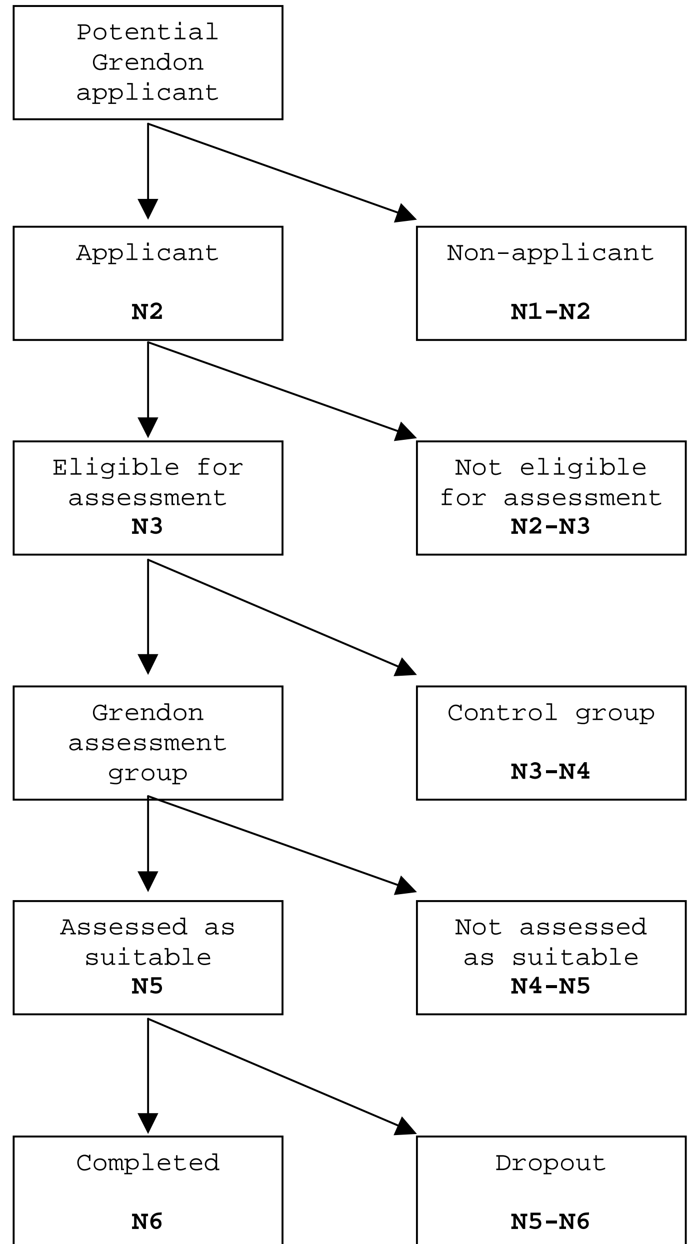
At present, N2 is approximately 140. (115 who enter and 25 on the waiting list). This would need to be increased considerably for an RCT to be viable.

Currently, very few are deemed ineligible at the application stage. An estimate of 10% dropout at this stage is therefore considered appropriate.

The statistical power analysis indicated that there needed to be at least 64 people in the treatment group. There should also be an equivalent number in the control group.

It is likely that around 75% of participants would be assessed as suitable following formal assessment.

Currently it is judged that only 44% of those treated at Grendon either partially or fully meet their therapy aims. Therefore a 56% dropout rate is assumed at this stage.



[6] The percentage dropping out of the study estimated here are the 'worst-case' scenarios. In practice, it would be important to try to minimise dropout at stages N5 and N6.

The largest sample scenario

Although the previous discussion gave valid calculations, it may be considered that only having 33 people successfully complete is too small a number to detect statistically significant differences. The importance of this is that if the effect size is small, then it is likely that the null-effect of the 67 people who dropped-out will swamp the improvement of the treatment group. An alternative to try to combat this is to ensure that there is a large number who successfully complete treatment. This would ensure that there is more chance of the effect size managing to make a difference to the overall results in this group.

Working back from N6, therefore, if there were 100 people successfully completing the Grendon treatment, there would need to be 303 people in each of the two groups, and therefore at least 673 applicants for places (Table 3.3).

Table 3.3: The number required with various scenarios

		Smallest sample scenario	Average sample scenario	Largest sample scenario
N2	The number of applicants required	222	444	673
N3	The number of applicants eligible for assessment (90% of N2)	200	400	606
	*** **RANDOMISATION*****			
N4	The experimental and control group sizes	100	200	303
N5	The number assessed as suitable for Grendon (75% of N4)	75	150	227
N6	the number successfully completing Grendon (44% of N5)	33	66	100

The average sample scenario

The numbers involved in the previous discussion are prohibitive. Although publicity and roadshows may be able to increase the number of people applying to Grendon, it is unlikely that the numbers will rise to such an extent that there will be close to 700 applicants who are all willing to participate in the study. In addition, as the numbers in the study increase, so does the work and expense involved: all of the people in the control group need to be followed up wherever they are in the prison system. Numbers of over 300 in the control group would be very difficult to follow up and keep track of over time.

A compromise, therefore, is to increase the numbers involved in the study, but not to such a degree that the implementation of the design becomes impractical. If the numbers in the first scenario are simply doubled, then a more realistic design emerges. This would involve 200 participants in each group, with a total number of applicants of 444. A respectable 66 individuals would be predicted to successfully complete therapy at Grendon and be in N6. This scenario would also overcome some of the problems associated with analysing sub-groups. With 200 people in the experimental group, there should be sufficient power to detect change in both the high PCL-R scoring population (currently 47 per cent of Grendon prisoners) and the lifer population (currently 50 per cent of the population).

Volunteers

The difficulty remains that, at present, the number of prisoners volunteering to go to Grendon is too small a sample from which to randomise. Currently, the authorities at Grendon have no problem ensuring that they have a steady entry and exit, and no difficulty filling beds. There is therefore a consensus that, although it would be nice to increase the waiting list, it is not a top priority. However, in order to run a successful RCT, it is essential for the waiting list to be

increased dramatically. This would need support from all tiers of staff within HMP Grendon and support from other parts of the Prison Service. It would also require resources.

Coupled with the difficulties associated with encouraging interest and applications to Grendon, a large obstacle to overcome will be ensuring that those on the waiting list agree to participate and be randomly assigned to either treatment or treatment as usual. With many other trials, notably those underway in the UK at present, randomisation is either to treatment or a waiting list control. That is, all participants will receive treatment at one time or another, but some will spend time on the waiting list while the first group is treated. Comparisons are made at the end of the first group's treatment, before the second group starts. In this way no one is denied treatment.

However, interviews with those working with the Grendon population revealed that they are quite confident in their short-term outcomes, and are confident that Grendon is successful at reducing the risk factors they think are associated with future offending. For them, the important question yet to be conclusively answered is whether Grendon has long-term effects, especially on future re-offending. In order to obtain this sort of comparison, it would be important to ensure that none in the control group was treated at Grendon before their release from prison. Although this may be justifiable ethically if only the criminogenic outcomes are considered (there is no clear evidence that Grendon is effective in reducing recidivism) it may not be justifiable on a personal level in that the study may be denying participants quality of life benefits.

Another difficulty is that, if a prisoner is assigned to the control group in this study, there is no reason for him not to then apply for a place at HMP Dovegate or some other TC. This would cause difficulty because the regime at HMP Dovegate more closely resembles that of Grendon than a more 'typical' facility in the Prison Service.

Finally, it is also possible that prisoners who have volunteered for a place at Grendon but are refused on the basis of chance will decompensate. That is, through disappointment or feelings of inequality or injustice their behaviour becomes worse than if they had never entered the study. This is particularly important with this group, as there may be many with paranoid personality disorder applying for places in Grendon.

A possible way of overcoming these issues is discussed in the section on 'the control group' below.

The control group

In order to protect against those in the control group decompensating, an attractive alternative could be developed as an alternative to Grendon. Understandably, an ethics committee may also find the idea of permanently prohibiting those in the control group access to Grendon problematic. To overcome this obstacle, the alternative to Grendon may be access to a personal advisor to help them through their prison life, or some other 'treatment' that may be seen as attractive and a reason to participate in the study.

In many RCTs the control group is largely overlooked. However, what happens to them is critical to understand whether there has been real benefit from the Grendon experience. It is important that all pre-treatment tests conducted with the Grendon group are also conducted with the control group. In addition, there are a good many treatments that those in the control group may be given over the course of the study; this too needs closely monitored.

It is also essential for those in both the experimental and control groups to fully understand the study and to participate in the full knowledge of what they were getting themselves involved in. This would involve detailed well-structured documentation on the study sent out with all blank application forms, and for the study to be discussed at the roadshows.

Timeframes

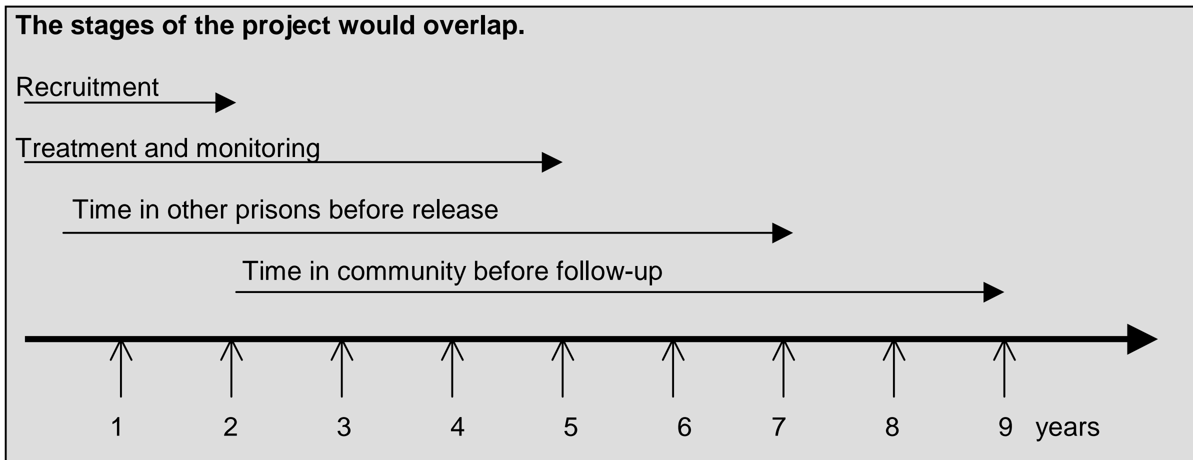
Crucially, many of the issues discussed above will have an impact on the timeframe of a possible RCT. Based on the 'average' case flow scenario, the study would require at least 200 people in each group and at least 444 applicants in the first instance (stage N1). In practice, an average of 115 prisoners enter Grendon every year. It would therefore take two years to ensure that the appropriate number of volunteers had entered the study (assuming the appropriate number of applicants could be recruited).

Once all the treatment and control groups are recruited, they would then need to be followed up for the duration of their stay at Grendon, typically between one and two years, but can be for as long as five. However, as stated earlier, it is likely that a large proportion of the treatment group will drop out of Grendon early, before they complete their therapy goals. Thus, any RCT would need to make pragmatic decisions about the length of follow-up of both the Grendon and the control group. A follow-up of three years may be the most likely to effectively measure the treatment gain in terms of changes to psychometric tests (previous research has demonstrated that 95 per cent of prisoners leave within three years).

This would not be the end of the evaluation: the most crucial element of an RCT at Grendon will be to compare the long-term reconviction rates of both groups. This will undoubtedly be a lengthy process, especially considering the high proportion of lifers in Grendon. The RCT would therefore have to allow time for a large majority of the offenders in the study to be released and for them to be at large. Traditionally, reconviction studies allow for at least two years at large, although with this population and the nature of their previous offences, the study would be more robust the longer the reconviction study was left.

The time course of a possible RCT would therefore look like this:

- Recruitment into the study 2 years
- Treatment and monitoring 3 years
- Time spent in other prisons before release 2 years
- Time spent in the community before follow-up 2 years



Thus, it is estimated that the RCT would take nine years from the start to the first reconviction study. However, to look only at the final reconviction figures at the end of nine years would be to ignore the rich data that would have been collected over that time, and the numerous interim and proxy measures, such as psychological change and adjudications, that can be monitored and measured. In addition, the data from this study can be followed up indefinitely, so that re-offending over the long-term can be compared with immediate reconviction and other measures.

Summary

For the study to be successful, it would have to be an integral part of the admissions process to Grendon for the duration of the project. Ideally, over the study period all entries to Grendon would also be entries to the study. This would involve wholesale enthusiasm for the project on the part of the governors, the therapists, and the prison officers at Grendon. In addition, there would also need to be support from beyond Grendon – Prison Service HQ and other prisons where applications to Grendon would come from.

At the start of this section, a number of obstacles to a successful RCT were raised. The discussion in this chapter has illustrated how these can be overcome. In sum:

- There would need to be support from a number of difference agencies including the Prison Service and all interested parties in Grendon.

  There is a great deal of support already for an RCT at Grendon. Through education and discussion and through working through many of the other potential obstacles, this enthusiasm could increase.

- Conducting an RCT would interfere with normal processes. It would necessitate redesigning allocation and selection.

  The design proposed here suggests that randomisation would occur before anyone enters Grendon. Thus, from the perspective of Grendon, the admissions process would not change at all. It is acknowledged that there would be a number of additional measures used to monitor progress and outcomes. However, this is no different to any other evaluation that may be used with the Grendon population.

- Fundamental to an RCT being feasible is increasing the demand for places – there are currently too few people on the waiting list.

  Previous attempts to increase applications for Grendon using roadshows have been successful. There is therefore no reason to suppose that it will not be successful again. In addition Prison Service support would help increase the numbers applying.

- Because of the pressure on beds in the wider Prison Service, it is thought that even if there were sufficient numbers on the waiting list, they would be given beds in Grendon as soon as they became available.

  If initiatives to increase the demand for places at Grendon are successful, there should be no reason to alter the rate of admissions from its present rate.

- If some in the control group are, in fact, DSPD it is likely that they will come into contact and be eligible for DSPD services when they come on stream in the future. This would severely affect the study in that it would no longer be possible to assert that the control group received 'treatment a usual'.

  This is an issue we cannot predict: we do not know how many of the group are currently DSPD or how many would be drawn into emerging services. However, there are contingency actions that can be taken as a last resort if such a problem occurred.

- At present 40 per cent of those on assessment and 50 per cent of those in the wings are lifers (they tend to stay longer in Grendon). If one of the most important outcome measures for Grendon is reconviction, then this may strongly affect outcome.

  It is possible to randomise in blocks to ensure equal numbers of lifers in the treatment and in the control group.

4. Conclusions

This examination of the feasibility of conducting an RCT to examine the long-term effects of the regime at HMP Grendon has revealed that it is indeed possible, although, like any large-scale evaluation, there are a number of obstacles that would need overcome first.

The primary obstacle identified was ensuring the support of all those involved in Grendon and within the broader Prison Service. However, this is easily answered: the benefits to Grendon of such an evaluation, both in terms of raising its profile and increasing its evidence base, are uncontroversial. Equally, another important obstacle to the success of an RCT is the numbers currently applying. However again, this should be easily overcome: roadshows have already been shown to increase the number of applications Grendon receives and support from the Prison Service for this study should also help raise its profile and thereby increase the numbers applying.

For an RCT to be successful, it needs to disrupt as few people as possible, and interfere as little as possible in the regime it is evaluating. For this reason, it would be most practical to randomise as early in the process as possible. The report suggests that this is ideally at the stage an application is received and a decision has been made that the applicant should be invited to Grendon for assessment. In this way, the randomisation process would not interfere with the normal admissions process at all.

However, there is a downside to early randomisation: it is likely that a large number of volunteers will drop out of Grendon during assessment and treatment for one reason or another. Previous work has shown that a number of prisoners are deemed not suitable for therapy, others infringe rules, and others are democratically voted out because of some on-wing behaviour. The report's 'worst-case scenario' estimate was that it is possible for only 33 per cent of the treatment group to successfully complete their therapy. If this is the case then it is important to ensure that as large a sample as possible start the treatment. Above this, minimising dropout at stages N5 and N6 (assessment and treatment) would help ensure the most robust results possible.

These are all preliminary discussions. A number of staff changes are currently underway at Grendon: a new governor is due to arrive; there is a vacancy for a new director of therapy and there is a vacancy for a head of psychology. Such staff changes are bound to have a very strong influence on the practical feasibility of conducting such a study. However, the message from this report that an RCT is possible, and the long-term advantages for Grendon and the Prison Service in general far outweigh the potential difficulties.

5. References

Chiesa, M. and Fonagy, P. (2000). *Cassel personality disorder study: Methodology and treatment effects.* British Journal of Psychiatry, 176, 485-491.

Farrington, D. and Jolliffe, D.. (2002). *A feasibility study into using a randomised controlled trial to evaluate treatment pilots at HMP Whitemoor.* London: Home Office. www.homeoffice.gov.uk/rds/adhocpubs1.html.

Genders, E. and Player, E. (1995). *Grendon: A study of a Therapeutic Community.* London : Clarendon Press.

Gray, N. S., Snowden, R. J., Brown, A. and MacCulloch, M.J. (unpublished). *Prevalence of psychopathy and other measures of risk at HMP Grendon: An investigation of population statistics.* Report submitted to the Home Office, 2002.

Gunn, J., Robertson, G., Dell, S. and Way, C. (1978). *Psychiatric Aspects of Imprisonment.* London: Academic Press.

Hare, R. D. (1991). *The Psychopathy Checklist - Revised.* Toronto: Multi-Health Systems.

Hobson, J., Shine, J. and Roberts, R. (2000). How do psychopaths behave in a prison therapeutic community? *Psychology, Crime & Law,* 6(2), 139-154.

Hobson, J. and Shine J. (1998). Measurement of psychopathy in a UK prison population referred for long term psychotherapy. *British Journal of Criminology.*, 38(3), 504-515.

Lipsey, M. W. and Wilson, D. B., (1998). Effective intervention with serious juvenile offenders: A synthesis of research. In R. Loeber & D. P. Farrington (Eds.) *Serious and Violent Juvenile Offenders: Risk Factors and Successful Interventions*, 313-345. Thousand Oaks, California: Sage.

Losel. F. (1998). Treatment and management of psychopaths. In D.J. Cooke, A.E. Forth and R.D. Hare (Eds) *Psychopathy: Theory, research and implications for society.* Amsterdam: Kluwer.

Marshall, P. (1997). A reconviction study of HMP Grendon Therapeutic Community. Home Office Research Findings 53. London: Home Office.

Newton, M. (1998). Changes in measures of personality, hostility and locus of control during residence in a prison therapeutic community. *Legal and Criminological Psychology.* 3(2), 209-223.

Newton, M. (unpublished). A reconviction study in HMP Grendon. Report submitted to the Home Office, 2001.

Redondo, S., Sanchez-Meca, J. and Garrido, V. (1999). The influence of treatment programmes on the recidivism of juvenile and adult offenders: A European meta-analytic review. *Psychology, Crime and Law*, 5, 251-278.

Rice, M.E., Harris, G.T., and Cormier, C.A. (1992). An evaluation of a maximum-security therapeutic community for psychopaths and other mentally disordered offenders. *Law and Human Behaviour,* 4, 399-413.

Robertson, G. and Gunn, T. (1987). A ten-year follow-up of men discharged from Grendon Prison. *British Journal of Psychiatry,* 151, 674-678.

Shine, J. (2000). *A compilation of Grendon Research.* HMP Leyhill: PES.

Taylor R. (2000). A seven year reconviction study of HMP Grendon Therapeutic Community. Home Office Research Findings 115. London: Home Office.

Appendix A: Questions and answers

This section details a list of common questions or criticisms frequently made in relation to RCTs in social settings. The reason such queries are or are not valid will be noted, followed by potential ways these criticisms/difficulties can be overcome. It is hoped that this section will help field some of the more common questions about RCTs.

What does the participating organisation get out of it?

RCTs can be seen as a lot of disruption for very little gain, with the evaluators taking most of the credit, when it is the people working at Grendon that are actually treating the patients. It can seem a valid criticism that the staff are being asked to do extra work for little reward. However:

- an RCT gives the opportunity for staff to learn from their work in a way that is more difficulty using other methods
- taking part in an RCT would increase Grendon's international prominence
- there is the potential for the work at Grendon to feed into government policy
- there is a need to obtain unambiguous independently evaluated outcomes about the treatment at Grendon
- the participation in the RCT and the results from it should add to Grendon's credibility and help with the accreditation process
- positive results from an RCT would increase the likelihood of continued funding.

An RCT at Grendon would put more work onto an already over burdened staff

It is true that being involved in an evaluation involves more work, particularly in the gathering and collection of assessments. However:

- this work will be no greater than being involved in any other type of evaluation
- every effort will be made to minimise the burden of random assignment and data collection on the staff.

An RCT gives too narrow an outcome

It is true that an RCT can only answer a very simple question, 'is the treatment provided at HMP Grendon of benefit compared with the services provided in the Prison Service in general?', but not other questions such as: *how* Grendon is treating people; or *why* the treatment manages to change them. However:

- an RCT does not have to be the whole story. Other analyses and pieces of data collection can occur simultaneously.
- it is important to be aware of the question being asked, because at the end of the day, an RCT cannot answer the question, 'are the services provided of benefit at all?'
- it is an important question that is being answered: it is pivotal to know if Grendon is providing additional benefit compared to the rest of the Prison Service and if it is working equally for all groups.

It is unethical to withhold treatment from the control group

It would be unethical to withhold treatment from a population if the relative effectiveness of the treatment is already known and if the population to be randomised was already receiving treatment.

However, it is not unethical if:

- the evidence is equivocal. With Grendon, this is the case, especially when the evidence concerning the treatment of psychopathic offenders is so controversial. An unambiguous evidence base is needed.
- there are only limited places in the treatment. Historically there always was a waiting list for places at Grendon. If the waiting list was to increase then arguably the fairest way to allocate places is via random assignment.
- another, potentially equally effective treatment, is offered. Although the waiting list would be denied access to Grendon, this is not to say that they are being denied treatment, only that they are being denied *this* treatment. However, it is essential that the services given to the control group are known, after all this is what Grendon is being compared with.

There are not enough people wanting to attend Grendon to produce the desired numbers for randomisation

Demand at present may be artificially low because of a lack of awareness of Grendon. People may not know that they may be eligible for Grendon, or understand it well enough to want to apply.

- roadshows and other such techniques may help increase the demand
- there needs to be strong support from the Prison Service as a whole to support the RCT and encourage volunteers from across the service.

Random allocation will result in the 'wrong' population

There is a perception amongst some that allocating places via randomisation will result in a different population than normal. Those usually assigning places do so on the basis of who is eligible and who is treatable. This would be ruined if experimenters assigned places. However:

- this is only a problem if the experimenter does not understand how places are usually assigned. However, if there is co-operation between those running the treatment and the evaluators, a mutual understanding and a commitment by the evaluator to work with the experts and accommodate their opinions then there should be little difficulty. It is in the best interest of both groups for the population from which people are randomly allocated to match the usual treatment group.
- if it isn't clear what services are best for which people, randomisation is arguably the best way of finding this out.
- at present there is an arbitrary nature to the selection process – it depends on who applies and who is given information about Grendon. An RCT would simply involve increasing the number applying and then randomising those selected as suitable for Grendon on the basis of their paper application.

What if randomisation results in an 'odd' sample by chance?

Random assignment, by its nature, can result in two populations that may be quite different by chance and therefore not comparable. However:

- if there are important differences for the evaluation, such as ethnicity or psychopathy then randomisation can be conducted in blocks
- it is also possible to generate a number of random configurations that are not peculiar and then randomly pick one of these.

Randomised experiments suppose an oversimplified theory of cause and effect

Some argue that multiple causal factors cause change often in complex, inter-related ways and therefore RCTs are not answering the right question. However:

- RCTs were not designed to answer questions about complex causal interdependencies – their purpose is more narrow and practical, 'does the intervention cause change in the measured outcome?'. They were designed to pull apart what is normally confounded.
- RCTs can give a clear picture of whether the link between two variables is causal, unlike other forms of evaluation.

RCTs are not possible because the sample is too heterogeneous

Many consider RCTs unsuitable to environments such as Grendon because of the diversity in the characteristics of the prisoners being treated and the variability of the resulting therapy. However:

- although there are heterogeneities, there are also commonalities. RCTs can be designed to capitalise on general similarities and by minimising the differences.
- also, heterogeneity can be acceptable in RCTs – and it can allow researchers to determine if a given outcome is effective across the range of known factors. However, to do this, statistical power must be sufficient to detect the main effect of treatment, despite the heterogeneity.
- it can be argued that the best way of overcoming the heterogeneities in the sample is by randomisation. Otherwise there may be systematic differences between the treatment and the control group.

Random assignment is premature until a good theory of the programme and its mediating processes has been developed

There have been arguments that RCTs are premature because they are based on emerging theories; implementation quality varies between units assigned to the same intervention; and implementation is often unfaithful to the underlying theory. However:

- randomisation does not require well-specified programme theories, or good management, or standard implementation or treatments that are totally faithful to programme theory, even though these features would make the evaluation much easier.
- in fact, the complexity of the organisation leads to the need for randomised experimentation by protecting against bias.

Randomised experiments are not suited to complex organisations

The following beliefs are widespread: fundamental change requires a political consensus that is often elusive, given the differing opinions of different politicians; each unit is so complex and established in its structure and functions that it is difficult to implement any changes that will modify its central functions; and each organisation is so distinctive in its goals and culture that the same change attempt will result in variable responses. However:

- holding all these beliefs will result in no evaluation occurring at all.
- it is not the case that these difficulties are prohibitive. North American research has demonstrated an impressive number of studies that have engendered strong political support and have involved large studies ranging from individual institutions to state-wide RCTs.
- it must be kept in mind that an RCT will not answer all the questions about a specific intervention and a number of studies may need to be completed to understand whether the same intervention produces the same results in different contexts.

Produced by the Research Development and Statistics Directorate, Home Office

This document is available only in Adobe Portable Document Format (**PDF**) through the RDS website

Home Office
Research, Development and Statistics Directorate
Communication Development Unit
Room 275
50 Queen Anne's Gate
London SW1H 9AT

Tel: 020 7273 2084 (answerphone outside of office hours)
Fax: 020 7222 0211
Email: publications.rds@homeoffice.gsi.gov.uk

ISBN 1 84082 964 8
© Crown copyright 2003