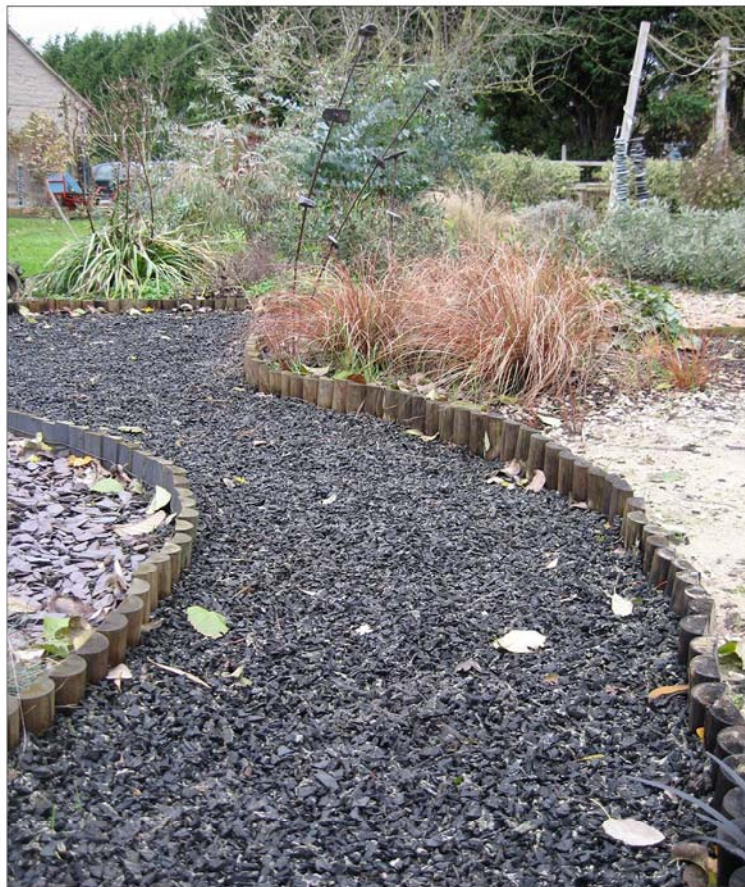


Researching social and therapeutic horticulture for people with mental ill health



a study of methodology

Joe Sempik

Researching Social and Therapeutic Horticulture for People with Mental Ill Health:

A study of methodology

Joe Sempik

Thrive, in association with the Centre for Child and Family Research,
Loughborough University.

First published in Great Britain in July 2007 by Thrive in association with the Centre for Child and Family Research, Loughborough University.

Thrive

The Geoffrey Udall Centre, Beech Hill, Reading RG7 2AT
Tel: 0118 9885688

www.thrive.org.uk

Centre for Child and Family Research

Department of Social Sciences, Loughborough University,
Leicestershire LE11 3TU
Tel: 01509 228355

www.ccfrr.org.uk

© Centre for Child and Family Research 2007

Joe Sempik (BSc, MSc, PhD) is a Research Fellow in the Centre for Child and Family Research, Loughborough University.

Acknowledgments

The author would like to acknowledge the Mental Health Foundation and Thrive for financial support of this study.

The author acknowledges the co-operation and collaboration of colleagues at Thrive, especially Tim Spurgeon and Allayne Amos. I would also like to thank all of the projects, their users, staff and volunteers; and the researchers and practitioners who took part in this study. Without their co-operation this work would not have been possible.

Citation

This report should be cited as:

Sempik, J. (2007) *Researching Social and Therapeutic Horticulture for people with Mental Ill Health: a study of methodology*, Reading: Thrive and Loughborough: Centre for Child and Family Research.

Cover – The Path of Life Garden at RF TWIGS, Swindon

The Path of Life garden was conceived and constructed by a group of people recovering from mental health problems. The concept is that of a path which is reflective of peoples' journey through their illness. The start of the path is bleak and barren, gradually softening and eventually developing into a serene and inspirational setting.

The garden is sometimes used as an evaluative tool, when individuals first join TWIGS they are invited to position themselves along the path at a point that represents their present stage of recovery. After a short period the same person would again be invited to repeat the exercise and any change is then reinforced by a combined visual, mental and physical process.

Alan Holland, RF TWIGS, Swindon, UK

Cover design by Peter Riley-Jordan, photograph by JS

Further information about this project is available on:
www.growingtogether.org.uk

Researching Social and Therapeutic Horticulture for People with Mental Ill Health: A study of methodology

Contents

1	Introduction	6
2	A Need for More Research? Views of researchers, practitioners and senior mental health professionals	10
3	Randomised Controlled Trials: a short history and their applicability to Social and Therapeutic Horticulture	20
4	Recent and Current Research: a view of three trials	38
5	Selection of Outcome Measures: results of a pilot study	44
6	Assembling the Sample	76
7	Inclusion and Exclusion Criteria	82
8	Defining the Intervention and Control Activities	84
9	Randomisation	96
10	Conclusions and Recommendations: a protocol for a trial of social and therapeutic horticulture	100
11	References	108
	Appendix	116
	Outcome Measures	116
	CORE	118
	HADS	122
	General Self-efficacy	126
	Use of Services	130

1 Introduction

Horticulture has been used as form of therapy or a complement to therapy for many years. It is generally practised as an organised activity referred to as a 'project'. The terms 'horticultural therapy', 'therapeutic horticulture' and more recently 'social and therapeutic horticulture' (STH) are usually applied to this practice. These are discussed in our review of the literature (see Sempik *et al*, 2003, p3) and also later in this report in the context of defining STH as an intervention for a study. STH is used by many different vulnerable groups and individuals, for example, those with mental health problems, learning difficulties, physical disabilities and other health and social problems (see Sempik *et al*, 2005).

Although the practice of STH and other nature-base approaches to promote health and well-being has been around for many years, the actual evidence as to its effectiveness is scant (see Sempik *et al*, 2003). There have been many articles written on the subject, but these have tended to be descriptive or instructional rather than providing empirical evidence as to effectiveness. However, we have recently completed a three year study of STH in the UK (known as the Growing Together study; see Sempik *et al*, 2005) in which we examined the benefits of STH against a framework of social inclusion proposed by Burchardt *et al* (2002). This framework has four key dimensions - *production, consumption, social interaction and political engagement* (see Sempik *et al*, 2005). We found that the benefits resulting from participation in STH projects were enacted through these dimensions. For example, garden projects provided the opportunity for social interaction and for some participants they were the *only* points of social contact. They enabled them to be *productive* through a variety of processes, for example, project participants were able to take part in activities which had many of the attributes of employment, such as purpose, routine and structure. Participants were engaged in meaningful activities which gave them the status of 'workers' or 'gardeners' and this promoted self-confidence, self-esteem and well-being. The research also showed that STH was an important and expanding area of

health and social care with around 1,000 active garden projects in the UK providing services for approximately 22,000 clients each week.

The main part of the Growing Together study was a qualitative examination, through interviews and observation, of the perceptions, thoughts and feelings of those who participate in STH projects, either as clients or as project organisers or volunteers. The study provided evidence as to *why* participants perceived STH to be beneficial; it provided in-depth information about the processes and mechanisms involved. But it could not test the hypothesis that STH was indeed effective. It could be argued that since (some of) the mechanisms and processes have been elicited, STH *must* therefore be effective. However, such an argument may be difficult to sustain in the context of increasing methodological rigour in the field of research into nature-based approaches for improving health and well-being. The present research, therefore, examines the methodological and practical issues associated with quantitative studies of effectiveness of STH and explores whether methods such as randomised controlled trials (RCTs), commonly used in 'mainstream' biomedical research, could be applied to STH.

Although STH is used for clients with a variety of psychological, physical and social problems, those with mental ill health comprise one of the largest groups of users (see Sempik *et al*, 2005). Therefore, this study has focused on that group. The involvement of that group may also facilitate recruitment into any future study, as this has been a problem in previous quantitative studies of therapeutic horticulture.

This report briefly examines the history of RCTs and their relevance to STH; it examines whether there is a real need for more research in the field and whether it would be possible to recruit sufficient participants in order to conduct a controlled study; it looks at three recent trials of nature-based approaches for health (allotment gardening, animal-assisted therapy and care farming) and the lessons that can be learned from them; it looks at potential outcome measures and inclusion and exclusion criteria; it seeks to describe a 'paradigm of social and therapeutic horticulture' which can be used to define

STH as an intervention in a study. Finally, it proposes a protocol for a study of STH which is feasible under current circumstances and with the resources that are likely to be available.

2 A Need for More Research?

Views of researchers, practitioners and senior mental health professionals

In our literature review of STH (Sempik *et al*, 2003) we highlighted the need for more research evidence and noted that whilst there are many published descriptive accounts of therapeutic horticulture and also some qualitative studies there was a distinct lack of quantitative data or experimental studies. A number of other researchers working in the field of health and horticulture research have also argued for controlled, quantitative studies. For example, Frumkin (2004) wrote:

We need a clinical epidemiology of horticulture.

What would this research look like? It would study well-defined populations with specific well-defined health conditions, recruited in large numbers to achieve a high level of statistical power. It would use randomised controlled trials whenever possible, to study well-defined clinical interventions.

...This research would assess health outcomes using accurate, reproducible, validated measures. It would control the bias and confounding that would otherwise prevent clear conclusions (Frumkin, 2004, p 23).

He also expressed similar views regarding the evidence base underpinning claims that experience of trees and woodland promotes health and well-being (Frumkin, 2005).

Whilst it appears that there is a perceived need in some quarters for more research and more results from experimental studies to support STH (and other nature-based therapeutic approaches), it is important to examine this assumption. If the view of a lack of evidence is only shared by a minority of researchers (who could be considered to be promoting their own interests) and the majority of researchers, practitioners and policy makers are satisfied

with the current evidence then more research does not necessarily address the needs of the STH community. In order to explore the apparent need for more research, opinions were sought from practitioners of STH; researchers working in the field of STH and associated therapeutic approaches; and senior mental health workers (consultant psychiatrists and senior academics working in mental health).

Practitioners

One hundred and nineteen practitioners of STH responded to an Internet-based survey on views relating to professionalisation of the practice, included in this survey were a number of questions on their perceptions of the current research evidence and a field for free text answers.

The sample of respondents represented a group of experienced practitioners. Seventy one per cent had been in practice for five years or more; 49% had been practising for 10 years or longer, and 27% had a formal qualification in therapeutic horticulture.

Only 21% of respondents considered the evidence base to be 'very strong' and 35% thought it to be 'strong', the remainder (45%) perceived it as either 'not so strong' or 'weak' (Table 1.1). When asked whether there were sufficient published accounts of quantitative studies only 13% agreed (see Table 1.2), however, 36% considered there to be sufficient descriptive studies. While these results suggest a perceived need for more research in general, it is clear that specifically the need is for more quantitative data. This was succinctly put by one respondent to the survey:

“Much more robust, quantitative research, and studies specific to single client groups [is needed]. There is a glut of qualitative and descriptive research and it is not well received by the more established health professionals.”

Table 1.1: Practitioners' views of the current evidence relating to Social and Therapeutic Horticulture

<i>View of current evidence base</i>	Frequency	Percent
Very strong	24	20.7
Strong	40	34.5
Not so strong	39	33.6
Weak	13	11.2
Total	116	100

Table 1.2: Practitioners' views of sufficiency of research evidence from different approaches (n=119)

<i>Are there sufficient published accounts of:</i>	Number responding 'Yes'	Percent
Descriptive studies	43	36.1
Case studies	33	27.7
Qualitative studies	33	27.7
Studies involving specific disability groups	23	19.3
Quantitative studies	15	12.6

Researchers

A group of international researchers active in the area of therapeutic horticulture and/or green care (use of land-based approaches to promote health and well-being of vulnerable people) who were attending a research workshop were asked to complete a short questionnaire regarding their views of the evidence base relating to STH and the current need for more experimental data and randomised controlled trials. Fourteen respondents returned valid questionnaires.

None of the researchers considered the evidence to be 'very strong' and only 21.4% considered it to be 'strong' (see Table 1.3). The majority (78.5%) rated it as 'not so strong' through to 'very weak' (compared with 45% of practitioners – see above). Nearly all of the researchers felt that there were sufficient published accounts of descriptive studies (93%) but none thought that there was sufficient quantitative research (see Table 1.4). Researchers, therefore, appear to perceive the lack of supporting evidence for STH more strongly than practitioners in the area.

When asked to rate the importance (on a scale of 1 to 10) of carrying out experimental studies and RCTs of STH the respondents returned a mean value of 8.29 (SD 1.978) for experimental studies and 7.14 (SD 2.568) for RCTs. Thus the overall consensus appears to be in favour of experimental studies and RCTs, but the view is by no means unanimous, as the following two contrasting quotations show:

“[We] will not be taken seriously in health services without this evidence [RCTs]”

“All HT gardens I know of are too small for randomised controlled studies. Alternative, trying to have studies on many gardens, you must be sure of having the same HT programs, standards, qualities in the HT gardens”

Table 1.3: Researchers' views of the current evidence relating to Social and Therapeutic Horticulture

<i>View of current evidence base</i>	Frequency	Percent
Very strong	0	0
Strong	3	21.4
Not so strong	8	57.1
Weak	2	14.3
Very weak	1	7.1
Total	14	100.0

Table 1.4: Researchers' views of sufficiency of research evidence from different approaches (n=14)

<i>Are there sufficient published accounts of:</i>	Number responding 'Yes'	Percent
Descriptive studies	13	92.9
Case studies	8	57.1
Qualitative studies	7	50.0
Studies involving specific disability groups	4	28.6
Quantitative studies	0	0

Senior Mental Health Workers

Discussions were held with a group of eight senior mental health professionals, which included four consultant psychiatrists (two of whom held academic chairs), the director of an institute of mental health, a university based psychologist who practised as a psychotherapist and two experienced researchers and mental health workers. They were asked for their views regarding the need for a randomised controlled trial of STH and then also asked to rate their perception, on a scale of 1 to 10, of the importance of conducting a trial at the present time. The scores ranged from zero to 10 (see Table 1.5). Although only seven participants provided a value, the mean was 6.29¹, slightly lower than that for researchers. However, this suggests the overall view is in favour of an RCT approach, although not without reservations.

¹ A value of '1' was used in place of the '0' given, as this was outside the range offered to respondents.

Table 1.5: Perceptions by senior mental health workers of the need for a randomised controlled trial of STH.

Respondent	Score	Comment
<i>a</i>	-	Not possible to score
<i>b</i>	0	No clear intervention, no clear target client group, no clear outcome measure therefore any RCT will have zero impact!
<i>c</i>	4	No design will answer all research questions...the RCT design is based on drug research, while psychosocial treatments have been subject to constraints for a number of decades. Such constraints include purity of intervention, homogeneity of samples, sample size, therapist effect, manualisation, comparison/control treatment, and controlling for confounders.
<i>d</i>	4	I think the idea of planning or being seen to plan, and seeking funding for an RCT is actually important. However the findings of any RCT are a double-edged sword (they can be sometimes be used to show that such approaches are ineffective) and unlikely to persuade anyone (they want replications, and nit pick on methods and generalisability).
<i>e</i>	7	Yes because, everybody's doing it and it's the only way to get recognised/noticed/respected/taken seriously, laws of the jungle and all that. A slightly quieter no because it's so methodologically arguable, so much more interesting and useful data could be gathered from the same effort by other methods, and RCTs are part of a pervasive top-down control freakery and arid reductionism that turns my blood to ice (and makes it boil!). But I must learn to live with it.

<i>f</i>	8	I think the only question that the Department of Health or Medical Research Council would be interested in is whether therapeutic horticulture engaged people with mental health problems that are difficult to engage and whether engagement leads to other health and social gains.
<i>g</i>	10	Politically I think all treatments at the moment need at least to try and get RCT based evidence, particularly if you want continued funding, so in terms of importance I would score 10. However, as we have discussed, the possible problems to be encountered may make it difficult to get findings that are going to be accepted, but I don't think this means we shouldn't try.
<i>h</i>	10	The problem is to specify the intervention and the outputs/outcomes clearly enough. If this can be done, there is no doubt that an RCT carries a lot of weight in policy terms. So ... if the practical/validity problems could be overcome, then I would go for a number nearer 10.

Three major themes were present in the comments; the need for RCTs to provide legitimacy among policy makers; the methodological limitations of RCTs for investigating psychosocial interventions and the need/difficulty to define the intervention or active component of STH. There was recognition that policy makers required RCTs, although a number of the respondents suggested that an RCT would not have a great impact because of “nit pick[ing] on methods and generalisability (respondent d)”. What was absent in the comments was any direct endorsement of the RCT method *per se* as means of collecting high quality data in the context of STH or psychosocial interventions. But instead an appeal to proceed with caution:

Overall, my view would be: consider an RCT design if the opportunity arises ... but do not put too many eggs in that basket. In other words, if you actively pursue an application to a range of bodies [for funding research], I would consider a range of designs to address different research questions. A qualitative design is a broadly obvious choice, but even this would need careful consideration on questions, sampling, measures, outcomes etc (i.e. not a straightforward either/or).
(Respondent c)

Similarly, respondent a, compared carrying out an RCT to betting on the favourite in a horse race – “you put a lot of money down and get little back”, whilst using qualitative methods was more like betting on an outsider where there was the potential of a large gain from a small outlay. In other words, the expense in carrying out an RCT may not be justified by the impact of such an approach.

Although some respondents considered that defining and replicating the intervention would pose problems to the successful execution of an RCT, others thought that these issues could be overcome, for example:

I don't see why identifying some of the processes involved in working out how therapeutic horticulture was effective in this context should be particularly difficult to establish. I agree that it would be essential to

establish what are the essential features of the intervention e.g. how much support, instruction and monitoring do they need, how much do addiction issues need to be also discussed, how much chasing up of people who attend poorly otherwise the intervention could not be successfully replicated and the generalisability and usefulness of the research would be limited. (Respondent f)

Conclusion

There is a general perceived need for more research, especially to generate quantitative data, among researchers and practitioners of STH and also among senior mental health professionals. Researchers, who might be expected to be familiar with the literature, perceived the lack of evidence more strongly than did practitioners and rated more highly the need for experimental studies and RCTs. Senior mental health professionals recognised the apparent importance of RCTs for policy makers and funding, but were concerned that data from RCTs of STH could be challenged on grounds of methodology and generalisability. They acknowledged the limitations of RCTs within psychosocial interventions and also placed a high value on qualitative data in that field. There is therefore a view in favour of RCTs but not without reservation.

3 Randomised Controlled Trials: A short history and their applicability to social and therapeutic horticulture

Randomised controlled trials are highly regarded within the field of medical research because they are considered a 'fair test' of a treatment. The importance of such fair tests is presented succinctly by the Editors of the James Lind Library². In their introduction to a series of essays on the evaluation of medical treatments, they write:

Misleading claims about treatments are common, so all of us need to be equipped to judge whether claims about the effects of treatments are valid. Without this knowledge, we risk concluding that useless treatments are helpful, or that helpful treatments are useless. Fair tests of treatment are tests that take steps to obtain reliable information about treatment effects by reducing the misleading influences of biases and the play of chance. When the need for fair tests of treatments is ignored, people suffer and die unnecessarily.

(Editorial commentary, The James Lind Library, 2007)

In order to carry out a fair test it is essential to reduce, or if possible to eliminate, bias in the selection of people and in their allocation to control and treated groups; and to ensure that any bias in the assessment of treatments is removed. At present the randomised controlled trial (RCT) is accepted to be a fair test of an intervention and occasionally has been referred to as a 'gold standard'. This is by no means an uncontested title and (as we shall see) there have been a number of criticisms and critiques of this approach and its appropriateness to certain situations has been questioned.

The RCT method involves the comparison of two (or more) treatments or interventions under conditions that remove any bias, either in selection of

² for a series of explanatory essays and a bibliography of the history of 'fair tests' of treatments in medicine refer to the James Lind Library, online at www.jameslindlibrary.org

participants or in the measurement of the outcomes of intervention, so that a *fair test* or comparison is performed.

The RCT was not developed as a single method within the medical research community; instead it evolved within a number of disciplines (medicine, psychology, education and social sciences). The development of the components advanced at a different pace within those disciplines and eventually came together to form the RCT that is used by medical researchers today. Whilst medicine accepted the RCT method, social scientists, who pioneered it, largely abandoned it in favour of other methods after the 1980s (see Oakley, 1998). In examining the conditions and issues relating to an RCT of STH it is useful to look at the history of the method in general, and to explore the important intermediate steps in the process. Whilst it may be possible to take the existing methodology ‘off the peg’ and use it in the field of STH, it is more likely that the method will need adaptation and development and may require a number of intermediate steps. Examining the history provides the hindsight that should make it easier to implement the RCT in the context of social and therapeutic horticulture.

The RCT has three essential components – *comparison, randomisation and ‘blinding’*.

Comparisons between treatments have been practiced for centuries. The James Lind library charts the use of such comparisons from Biblical times to the present day, starting with a passage from the Book of Daniel in which Daniel and his friends refuse Nebuchadnezzar’s meat and wine in favour of a diet of pulses and water, and are then compared to those who take the king’s food:

“And at the end of ten days their countenances appeared fairer and fatter in flesh than all the children which did eat the portion of the king’s meat” (Daniel, Chapter 1, v. 15).

Whilst such a test is clearly not scientific, it illustrates the power of comparing two (or more) treatments or approaches. It is important, however, that appropriate and not misleading comparisons are made. History provides other, carefully documented and meticulously executed examples of comparisons of different treatments. In 1753 James Lind published his account of a comparison of six different treatments for scurvy (*A Treatise of the Scurvy*; see Tröhler, 2003) in 12 patients suffering from the disease at sea. What is particularly remarkable is that not only did he include his own, personal account of his experiment but he added a summary and evaluation of the known writings on scurvy at that time and accounts sent to him by members of the Society of Naval Surgeons. Such an approach today would be regarded as a *systematic review*. He started his work by rejecting previously held, unsubstantiated beliefs regarding scurvy and declared:

"Indeed, before the subject could be set in clear and proper light, it was necessary to remove a great deal of rubbish" (Lind 1753, p viii in the James Lind Library, 2007).

Although Lind only used two patients per treatment the effect of administering oranges and lemons was dramatic, however, fruit juices were not recognised by the Admiralty as preventing scurvy until the 1790s. Other, cheaper competing treatments were tested and promoted by their supporters. Lind's *Treatise* is now regarded as a historical milestone in clinical research but although he was highly regarded as a naval surgeon at the time and his work was carefully conducted it did not have an immediate impact. It is also worth mentioning that the effects seen by Lind were indeed dramatic, but involved a very small number of patients.

Although, historically, conclusions about treatment have been made on the basis of dramatic results obtained from small samples, for example, Banting and Best's discovery of insulin (Banting *et al*, 1922) or Lister's use of 'the antiseptic system' in the Glasgow Royal Infirmary (Lister, 1870) these are rare. And as mentioned above, even dramatic results can be slow at influencing policy, especially if there is a cost involved. The effects of many

treatments, however, are subtle and need carefully controlled conditions to distinguish them from background noise or the effects of chance.

Whilst Lind compared different treatments which had been considered to be effective (comparisons between active treatments are still carried out today), another approach was to compare the effects of a treatment with no treatment or with one that is known to be ineffective i.e. an inactive 'control' or 'placebo'. The term "control" entered the scientific language in the 1870s to mean a standard against which an experimental intervention could be compared. Its main use was in experimental psychology. The word itself is derived from "counter-roll", "a duplicate register or account made to verify an official account" (see Oakley, 1998, p 1240). Controlled studies were carried out in education, psychology and social policy and the 'experimental method' in those fields developed largely independently of that in medical research.

One important aspect of the experimental method is *randomisation*. When comparing the effects of two different treatments it is essential that the groups receiving those treatments are similar, i.e. that any factors that influence the outcome are evenly distributed between the groups. If they are not, it cannot be safely concluded that any differences between responses of the two groups are caused by differences in treatment. They could just as easily have been caused by differences in the makeup of the groups themselves.

The earliest reference to using a random allocation to compare different treatments or approaches is attributed to the Flemish physician, John Baptista van Helmont, who wrote the following in his book *Oriatrike, or Physick Refined* (translated and published in England in 1662)

"Let us take out of the hospitals, out of the Camps, or from elsewhere, 200, or 500 poor People, that have Fevers, Pleurisies etc. Let us divide them in halfes, let us cast lots, that one of them may fall to my share, and the other to yours; I will cure them without blood-letting and sensible evacuation; but do you do, as ye know (for neither do I tye you up to the boasting, or of Phlebotomy, or the abstinence from a solutive Medicine)

we shall see how many funerals both of us shall have: But let the reward of the contention or wager, be 300 florins, deposited in both sides: here your business is decided³"

(quoted also in part by Doll, 1998; Manning, 2004)

van Helmont had issued a challenge to the medical establishment of his day to justify the use of their remedies by comparing theirs (which were based on theory) with his, which were founded on his experience of practice. Such a comparison was never made, but randomisation eventually became an important principle in the design of a fair test or comparison of a medical intervention.

It is often quoted that the Medical Research Council trial of the antibiotic streptomycin in pulmonary tuberculosis, published in 1948 (MRC, 1948), is the first 'true' randomised controlled trial. This used individual randomisation in which patients were allocated sealed envelopes containing a piece of paper with either the letter 'C' for control or 'S' for streptomycin. This eliminated bias, as the doctors did not know to which group any individual would be assigned and could not therefore influence that decision. However, there are a number of other contenders for the title of 'first RCT' as many earlier studies had features which included a form of random allocation and comparison group. For example, the MRC whooping cough trial (a RCT with individual randomisation) started before the streptomycin trial but reported after its publication in 1951; Amberson's trial of sanocrysin in tuberculosis (Amberson, McMahon and Pinner, 1931) which used 24 patients in 12 matched pairs is also a contender. Many years before that, however, a Danish physician, Johannes A G Fibiger, allocated patients to different treatment groups according to their day of admission in his trial of diphtheria serum in 1898 (see Hróbjartsson *et al*, 1998). Allocation to groups was by alternate days.

3

http://www.jameslindlibrary.org/trial_records/17th_18th_Century/van_helmont/van_helmont_kp.html

The previous examples given are from the field of medical sciences, but as mentioned previously, the experimental method owes much to psychology, education and social sciences. Oakley (1998) writes that “the psychologist CS Peirce introduced both the idea of randomisation and that of ‘blindness’ (see later) into psychology experiments in the 1880s” (p 1240). She also quotes from William A McCall’s method of randomisation published in 1923 in his book *How to Experiment in Education*

“One method of equating by chance is to mix the names of the subjects to be used. Half may be drawn at random. This half will constitute one group while the other half will constitute the other group” (Oakley, 1998, p 1240).

Other methods of ‘random’ (or pseudorandom) allocation have been used in the past, for example, alternation – assigning alternate participants to a control group or treatment group. This method is susceptible to bias because the researcher is able to manipulate which patients are allocated to which group. For example, the researcher could exclude a subject if he or she suspected that by including that particular individual the results of the experiment would come out unfavourably. This is known as ‘allocation’ or ‘selection’ bias.

The proof that such selection bias does indeed occur lies in the observation that studies which do not use randomisation or concealment of allocation usually show greater effects, on average, than those that do (see Chalmers, 2003, p 28; Kunz and Oxman, 1998). Randomisation in clinical trials has become one of the quality criteria against which the published research is judged and even the methods by which randomisation is achieved come under close scrutiny (see Higgins and Green, 2006, p 81). It is important that there is ‘allocation concealment’, i.e. that researchers do not know (and hence cannot influence or bias) to which group or treatment a study participant will be assigned. Research has shown that inadequate concealment of allocation leads to bias in the results of a study (see Chalmers 1983; Schulz 1995; Mohler 1998 quoted in Higgins and Green, 2006, p81). In other words, the

researchers influence the allocation process (either deliberately or subconsciously) so that the study produces more favourable or desirable results. Randomisation, however, is not always possible because of practical, ethical or other reasons. Whilst the preceding paragraphs have presented the case for randomisation *in general*, the practical issues relating to randomisation in a trial of STH in the UK are discussed in Section 9 of this report.

'Blindness' or 'blinding' in randomised controlled trials refers to the procedure of not revealing the treatment to the participant (single blind); or also of not revealing it to the researcher administering the treatment or carrying out measurements or assessments (double blind). In a double blind trial, for example, neither the participant nor the researcher knows which treatment a participant is receiving as allocation is performed by separate, impartial individual or group. This reduces 'ascertainment' or 'detection' bias. The researcher or assessor cannot influence the results (through, for example, the measurement or data analysis process) since he or she does not know which treatment has been given. This is true in theory at least, since it is often possible for a researcher to deduce the treatment that an individual has been given by their response to it. One way to overcome this is for the assessments and analysis to be carried out by other, independent researchers who are also 'blinded' to the treatments given. Blinding, like randomisation, is one of the criteria of quality in clinical trials.

Blinding is sometimes confused with allocation concealment; or both terms taken to mean the same. They are separate procedures which safeguard different parts of the RCT protocol. Their respective functions have been described by Schulz (2000) as follows:

Allocation concealment should not be confused with blinding. Allocation concealment concentrates on preventing selection and confounding biases, safeguards the assignment sequence *before and until* allocation, and can always be successfully implemented. Blinding concentrates on preventing study personnel and participants from determining the group

to which participants have been assigned (which leads to ascertainment bias), [blinding] safeguards the sequence after allocation, and cannot always be implemented.

Like randomisation, blinding is not always possible. Not all treatments and procedures are amenable to blinding – for some interventions it is obvious that they have been administered. Such is the case for STH. Different approaches need to be taken to safeguard the integrity of a study and eliminate bias.

Objections to RCTs

Randomised controlled trials have been used not only in laboratory experiments in psychology and education, but also in social policy on a grand scale. In the US, in particular, many large scale social experiments were carried out involving thousands of participants on interventions such as housing allowances and job search assistance. The peak of this ‘golden age of experimentation’ took place between the 1960s and 1980s (see Oakley 1998; see also Dehue 2001 for a sociological perspective on experimental methods in social policy). Afterwards social scientists gradually abandoned the method for other approaches, initially reflecting a change in political and ethical attitudes rather than a dissatisfaction with the reliability or validity of the method itself. However, in time this appears to have changed into hostility against the RCT method by some social science researchers and also by some researchers working in the field of complementary and alternative medicine.

A number of different objections have been made against the RCT design. These objections have targeted the ethics or validity of the whole approach or have been lodged against individual components of the RCT, for example, randomisation. Others have chosen to criticise the reliance placed on evidence-based medicine by the medical establishment and its apparent rejection of observational data. For example, Smith and Pell (2003), in an article in *BMJ*, called for a randomised controlled trial of the parachute since

evidence of its effectiveness was based only on observational data. They wrote:

As with many interventions intended to prevent ill health, the effectiveness of parachutes has not been subjected to rigorous evaluation by using randomised controlled trials. Advocates of evidence based medicine have criticised the adoption of interventions evaluated by using only observational data. We think that everyone might benefit if the most radical protagonists of evidence based medicine organised and participated in a double blind, randomised, placebo controlled, crossover trial of the parachute. (Smith and Pell, 2003, p 1459)

Perhaps not surprisingly researchers have come out against RCTs in social policy because of ethical qualms, for example, the social researcher, F. Stuart Chapin argued that “experimental allocation to social measures collides with the humanitarian mores of reform” (see Dehue, 2001, p 292). Others have argued that withholding a treatment from a control group is unethical. Such an objection can only be relevant if the intervention has already been shown to be effective or beneficial using a fair test. However, an RCT is only appropriate when there is uncertainty as to the effectiveness of the intervention being tested. Also, an experimental design can be used, such as that with a ‘waiting list control’, which allows all the participants in a study to receive the intervention. Those in the control group receive it later than those in the treated group. A similar approach is the ‘stepped wedge’ design (see Brown and Lilford, 2006) which involves the “sequential roll-out of an intervention to participants over a number of time periods” (p. 1).

Some researchers from the social sciences reject the notion that RCTs provide results which are of high quality. Becker, Bryman and Sempik (2006), for example, reported that in a survey of social policy researchers only 32% of respondents associated high quality results with experimental designs such as RCTs. Some researchers considered the RCT design inappropriate for complex situations. They report one respondent to the survey who wrote “that

it was crucial that the research approach “appropriately addresses issues of complex causation. N.B. randomized controlled trials don’t!!!” (Becker *et al*, 2006, p11).

There has also been debate regarding the use of RCTs within the medical field, particularly within psychiatry. Many researchers have argued that the complex nature of interventions such as psychotherapy make it very difficult, if not impossible, to carry out adequately controlled trials using a standardised ‘dose’ of intervention. Manning (2004), for example, explored the potential of RCTs in researching the effectiveness of the therapeutic community approach to mental ill health. He concluded:

The RCT is for many observers of medical and social practice a powerful method of developing a strongly legitimate means for gathering evidence which carries extensive social power.

However, the RCT as practised is not an appropriate gold standard solution for all problems. It certainly cannot be the required standard for an assessment of the therapeutic community movement, or a single local therapeutic community. While it could answer some questions about therapeutic communities, there would be massive problems and large costs. This is not to say that RCTs should not be done where appropriate.

Other approaches may be needed first, though and continued monitoring of therapeutic communities through a variety of assessment methods will be necessary not only to replace RCTs if cost or feasibility rules them out, but also to check whether RCT results are sustainable and generalisable. (Manning, 2004, p 119)

One of the earliest arguments against randomisation and the ‘statistical testing’ of interventions was that it did not identify those individual patients who would benefit from treatment – instead it showed the collective effect on the whole group. This objection was voiced by Dr Lewis, the physician in

charge of the department of clinical research at University College Hospital in the 1930s:

“it is to be recognised that the statistical method of testing treatment is never more than a temporary expedient, and that but little progress can come from it directly: for in investigating cases collectively, it does not discriminate between cases that benefit and those that do not, and so fails to determine criteria by which we may know beforehand in any given case that treatment will be successful” (Lewis, 1934 quoted in Doll, 1998, p 1219).

Today, some social scientists still voice their criticisms of randomisation and display their apparent lack of experience of the method within the literature, see, for example, Chalmers (2003) who neatly summarises the implications of these objections:

“Those who reject randomization are implying they are sufficiently knowledgeable about the complexities of influences in the social world that they know how to take account of all potentially confounding factors of prognostic importance, including those that they have not measured, when comparing groups to estimate intervention effects” (Chalmers, 2003, p 30).

With regards to social and therapeutic horticulture, we may not be certain which particular individuals will benefit from attending a garden project but that is not the point of carrying out research. The purpose of the research is to investigate whether the intervention is actually capable of providing improvement in the health and well-being of those clients for whom it is considered to be an appropriate ‘treatment’ *and not necessarily of all those with a similar condition*. The applicability of the intervention to individuals (i.e. its indication) is a matter for practitioners and referrers. They need to know, however, that when used appropriately the treatment is effective. For that reason it is necessary to subject the intervention, social and therapeutic horticulture in this case, to a fair test, and that may involve applying it

'collectively' to a random sample or other group that is judged to be appropriate or practicable.

Assessing the quality of RCTs

Many of those who criticise RCTs point to the inadequacies and shortcomings of poorly conducted trials, however, RCTs will only provide high quality evidence if they properly carried out. Simply labelling a study as 'randomised' and 'controlled' does not automatically imbue it with quality. Indeed, many researchers in medicine have complained about the quality of RCTs in their discipline and this has been a long-standing area of research and debate (see, for example, Jadad, 1998). Essentially, quality in RCTs can be seen in terms of those features of the trial – the design and conduct - that minimise bias. A high quality clinical trial is one that has low bias and is therefore a fair test of an intervention. Sources of bias in an RCT have been discussed above but can be usefully summarised here:

- selection bias – the makeup of comparison groups is not the same;
countermeasure: randomisation, i.e. equal distribution of confounding variables between groups
- performance bias – there has been a systematic difference in the way that comparison groups have been treated apart from the intervention being tested
countermeasure: blinding of those providing and receiving interventions; blinding of assessors as for detection bias (see below)
- attrition bias – there has been a selective drop out from one of the groups so that by the end of the study there are differences between them that have not been caused by the intervention
countermeasure: adequate reporting of dropouts and withdrawals from study; use of 'intention to treat' analysis (see later)

- detection bias – there are differences in the way that assessments have been made of the comparison groups
countermeasure: blinding of assessor, those carrying out the assessment or measurement of effects do not know which treatments participants have received

(adapted and summarised from Higgins and Green (2006))

A number of different systems have been used in an attempt to score the quality of RCTs by examining the sources of bias listed above. One of the simplest and most used scales is that created by Jadad *et al* (1996). This calculates a total score, on a scale of 0 to 5, based on the reporting of randomisation, double-blinding and reporting of dropouts. See Box 3.1. It is important to note here that assessment of quality is based on the *reporting* of the trial in the literature and not necessarily on the quality (validity) of the trial itself. However, the quality of reporting also contributes to the quality of the trial *as a whole*, for example, by making the research process transparent and providing sufficient detail for the study to be evaluated or repeated by other researchers. Adequate reporting of all methods, results, dropouts and withdrawals contributes to the overall quality of a study. It has to be accepted (unless there is evidence from other sources) that the trial has been conducted as described in the account.

Box 3.1: Calculating the quality of a clinical trial using the Jadad Score

	Score 1 point for each 'yes' or 0 points for each 'no'.
1	Was the study described as randomized (this includes words such as randomly, random, and randomization)?
2	Was the study described as double blind?
3	Was there a description of withdrawals and dropouts?
	Add 1 additional point if
	For question 1, the method used to generate the sequence of randomization was described and it was appropriate (table of random numbers, computer-generated, etc)
and/or	If for question 2, the method of double blinding was described and it was appropriate (identical placebo, active placebo, dummy, etc)
	Deduct 1 point if:
	For question 1, the method used to generate the sequence of randomization was described and it was inappropriate (patients were allocated alternately, or according to date of birth, hospital number, etc).
and/or	For question 2, the study was described as double blind but the method of blinding was inappropriate (e.g., comparison of tablet vs. injection with no double dummy).

(Adapted from Jadad *et al*, 1996, pp 10-11)

Application of RCTs to Complex Interventions and to Complementary and Alternative Medicine

The RCT design is seen as a fair test and has been used successfully in many different applications, including medicine and social policy. RCTs are also applicable to complex situations but, as interventions become more complex and difficult to standardise (or when standardisation actually interferes with treatment), the experimental design becomes more challenging and more expensive, but that is not to say impossible. The easier and cheaper option is often to reject the method in favour of an alternative approach, for example, a qualitative method. However, that does not necessarily solve the problem as there is now a socio-political recognition of the power of the RCT and an expectation that *any*, and indeed *every* intervention should be subjected to this method of evaluation. Whilst some of those, for example, in the field of complementary and alternative medicine (CAM) have argued that RCTs are not applicable to their discipline and have even put forward highly complex and technical arguments to support their standpoint (see Milgrom 2005) the medical establishment and policy makers have continued to judge the effectiveness of CAM (and other complex interventions) by the quality standards of the RCT. For example, Linde *et al* (2001) assessed the quality of published research on CAM using the Jadad scale described above and found that the majority had important shortcomings in methodology and in their reporting of the trials. The major problems were with the description of allocation concealment and with their handling of drop-outs and withdrawals. They also reported an almost total lack of 'intention to treat analysis', i.e. the inclusion in statistical analysis of all participants in a study, even those that have dropped out (see Hollis and Campbell, 1999). In a commentary on their analysis Edzard Ernst, Professor of Complementary Medicine at the University of Exeter, UK, wrote:

While we scientists lament the low average quality of RCT, providers of CAM very often have quite a different agenda and rarely feel the need for scientific scrutiny at all. They often argue that CAM, for a number of reasons, defies the straight jacket of reductionistic science. Most

readers of the *International Journal of Epidemiology* will agree that this attitude must be based on misunderstandings...What is true, however, is that scientists are constantly and miserably failing to get their points across to advocates of CAM.

Randomized clinical trials of CAM are often more difficult and methodologically more challenging than RCT of other types of interventions. Due to the nature of most CAM modalities and the conditions they are used for, such RCT often need to be large, of long duration and require expensive therapists' time. In turn, this means that CAM research is expensive and requires high levels of expertise in terms of trial design (Ernst, 2001, p 532).

Evidence from high quality, well-conducted clinical trials is therefore expected for CAM and other 'therapeutic' interventions, including also social and therapeutic horticulture. Simply rejecting RCT methodology as 'inappropriate' is increasingly becoming less of an option. Any published research on STH will eventually be judged on the criteria described above and ultimately its role in health policy will be determined by the quality of evidence put forward in its support.

The Relevance to STH Research – what are the lessons of history?

Our brief look at the history of RCTs shows that they took many years to be developed into the tools that they are today. They have been refined to be procedures which provide reliable answers to questions of efficacy and effectiveness of interventions in medicine. They work best (and are easiest to apply) in situations where those interventions are clearly defined. We have seen that in areas of complex interventions, or those such as homeopathy which exist at the fringes of medicine, they are a contested issue. The socio-political power of RCTs to grant legitimacy to medical treatments is recognised and this has, in some cases, shaped the attitudes of researchers and practitioners to RCTs (and produced hostility). Yet RCTs have been

carried out successfully using many complex interventions in social policy. Criticisms of RCTs *in general* have sometimes been based on specific studies *in particular* that have been conducted badly, rather than on a critique of the method itself. From a scientific, theoretical or ethical point of view there is no reason why an RCT approach should not be appropriate for STH. Issues that need to be considered are those of practicality and of the cost-effectiveness i.e. whether the *impact* of such a study *on policy or practice* would be sufficient to warrant the costs of such research and whether an RCT of STH (even if it shows clear and positive results) is likely to affect health policy, funding in the sector or on referrals to STH projects. These issues are discussed later in the report.

From a *theoretical* point of view most of the necessary conditions for an RCT are possible in the case of social and therapeutic horticulture. These include comparison, randomisation (including allocation concealment) and blinding of researchers conducting analysis; but not of the participants undergoing the intervention. Whether all of these conditions can be fulfilled in a practical sense is discussed throughout this report.

4 Recent and Current Research

A view of three trials

Whilst there have been many calls for researchers to conduct controlled research in the field of nature based approaches to health, including therapeutic horticulture (see Section 2), there has been little progress in that area. There has, however, been a small number of trials. These can shed light on some of the difficulties associated with such research and also on strategies for dealing with those difficulties.

Milligan et al (2003) set out to investigate the effects of allotment gardening on the health and well-being of older people in Northern England. Their intention was to recruit 300 participants into a study comparing a group engaged in allotment gardening with a social club and a 'reference' group. Their chosen instrument was the SF36 health and well-being questionnaire (see, for example, Jenkinson *et al*, 1999) and although they do not provide details of their statistical power calculations or their expected effect sizes, it is likely that had they been able to recruit such a sample size their study would have had the required power. However, they were only able to recruit a sample of 93. The attrition rate in their gardening group was 45%; 24% in the social club and 21% in the reference group. This meant that at the end of the study there were only 16 participants left in the gardening group and 22 and 21 in the other two groups. Thus by the end, the study lacked the statistical power to detect any modest changes in outcome measure between the groups. This study highlights both the difficulties of recruiting and retaining an adequate sample and the consequences of failing to do so. It also shows the need to adopt a flexible approach to the methodology should the researcher suspect that recruitment could pose difficulties. Milligan et al (2003) included in their design a qualitative dimension to the study (using diaries, focus groups and interviews) which they effectively maximised when they encountered difficulties with the quantitative element.

Their method of recruitment had been to select a sample from the lists of general practitioners (GPs). However, they encountered delays in receiving lists of eligible participants from the GPs, who presumably were busy with other tasks. Milligan et al (2003) do not give details of the number of participating GPs, or the number that they approached, or the number that declined to be involved. Whilst health and other professionals (in many different disciplines) may show an initial enthusiasm for a study there are often competing pressures which ultimately lead to disengagement. Obtaining commitment to a study from the obvious 'gatekeepers' is often seen as the initial part of the research itself and time for this is allocated in research proposals. However, the success of the whole study rests on the success of this initial phase. But the 'clock is ticking' from the start for most studies which are required to be completed within a finite period. Any delay or extension in the recruitment period results in a concomitant shortening of the intervention, follow-up or analysis phase. Therefore, there is good reason to separate the preliminary commitment/recruitment phase from the main study and allow sufficient time for securing the commitment of gatekeeper agencies and individuals. There is then also the opportunity to try another approach should a first one fail. For example, one approach that we had considered for a randomised trial of STH was for random allocation to control and intervention groups to be carried out by community mental health trusts (CMHTs), yet when we approached the trusts there was little interest in such involvement and an alternative strategy has had to be considered (see Section 9).

Another study of relevance is that of animal assisted therapy in Norway, conducted by Bente Berget (Berget, 2006). This was a randomised controlled trial which was completed successfully. It involved an initial sample of 90 psychiatric patients, randomised to a 'farm' group (60 patients) and a control group (30 patients). The attrition rate was again higher in the intervention group (32%) compared with the control (7%). The researchers had no apparent difficulty in recruiting the patients to the study through six participating institutions which also carried out the random allocation. They successfully used multiple centres for the research. Fifteen farms were recruited to facilitate the contact between patients and animals, i.e. to provide

the intervention. The participants had a variety of different diagnosed psychiatric conditions and this presented no specific difficulties to the conduct of the study or to the analysis of the data. However, because of the relatively small sample (and small numbers within each diagnostic category) any inference relating to the effectiveness of animal assisted therapy for different conditions should be treated with care.

Berget's work showed that the effect size with such an intervention is likely to be small (around 0.5 Standard Deviations) and that an effect may not be seen with all instruments measuring mental health and well-being. She also demonstrated the potential importance of follow up assessments after completion of the intervention. In her study the differences in measures of anxiety and general self efficacy were observed six months after completion of animal assisted therapy and were not apparent immediately after the intervention. However, follow up is not always possible, not only because of researchers losing touch with participants, but also because some study designs (see Section 10) recruit participants who have been referred to interventions (such as garden projects or care farms) without any limits on their stay. In these circumstances there would be serious ethical issues if participants were asked to leave projects to satisfy the needs of a research protocol.

We have also considered Marjolein Ellings's study of care farms in the Netherlands (personal communication). This three year study is currently in progress and due to complete in June 2008. It is a controlled study without randomisation, because of objections to this raised by the ethics committee. It aims to compare the perceived mental health and well being of a heterogeneous sample of people with mental ill health and history of substance misuse who attend care farms (intervention) with those who work in bicycle repair shops or woodworking workshops (control). The researchers reported that their power calculations showed that 150 participants would be required for each of the groups. However, by June 2006 a total of only 100 participants had been recruited into the study, and 25 had dropped out, leaving 50 in the intervention group and 25 in the control. The loss of

participants has been predominantly among those with substance misuse problems. Because of this attrition, and the slower than expected rate of recruitment, the recruitment phase of the study was extended and a final assessment, scheduled to take place at 18 months, has been cancelled.

Recruitment of participants into the study has been through a national institute of psychiatry, even so, the level has not been sufficient to reach the original target. Again, this illustrates the vulnerability of relatively large studies with a fixed time constraint to difficulties caused by slow recruitment of study participants.

Fifty care farms have also been recruited to supply the 'intervention' (compared with 15 in Berget's trial). Issues regarding the homogeneity of the intervention have not been raised by the researchers, however, they have expressed some concern that a number of participants were not maintaining the required level of attendance (two days per week). Issues of attendance i.e. the degree of intervention received, will be addressed by the researchers during the course of data analysis.

The researchers have also commented on the difficulties in administration and management presented by the use of such a large number of service providers. Some of these difficulties are concerned with the assessments. These are carried out by the care farmers. The researchers, however, have to keep track of the timing for each individual participant and inform the care farmers of impending assessments. This has caused problems because the researchers are also involved with a number of other projects that compete for their time and attention. In retrospect, the researchers would have preferred to have conducted a smaller trial in the first instance, for example, one that involved an in-depth examination of a smaller number of participants at fewer care farms.

The researchers have also included a qualitative element in their study to complement the quantitative data. This will explore participants' perceptions of the impact of the care farms on their health and well-being through interviews.

They will also follow up the participants who have dropped out of the trial in order to explore the reasons for their decisions to leave.

These accounts highlight the importance of transparency – providing detailed accounts of methods, results and difficulties so that the research can be judged; and so it can inform other studies, such as the one described in this report. For example, the description of the difficulties encountered by Milligan et al (2003) in recruiting a sample is likely to be useful to others contemplating a similar approach; the detailed results provided by Berget, including scores for general self efficacy, were useful for comparative purposes in this work (see Section 5). A greater use of common outcome measures would facilitate such comparisons of findings. However, national preferences for different outcome measures and language differences mean that this is not always possible, although some outcome measures, such as the World Health Organisation's quality of life instrument (WHOQOL-BREF) and the General Self Efficacy scale have been translated into many different languages and normative (and some clinical) data are available from a number of different countries.

Transparency is one of the key indicators of the quality of a trial (see section 2) and must be addressed in the reporting of any future trial of STH, whether successful or not. This will enable the development of a rigorous methodology for the study of STH and other nature-based complex interventions.

Summary and Recommendations

- Issues of recruitment and attrition pose a serious risk to a study of social and therapeutic horticulture (and other complex interventions based on nature-orientated approaches). It is essential that the recruitment strategy is well-thought out (and preferably tested) prior to commencement of a study.
- A preparatory/recruitment phase which is separate from the main part of a study may improve the likelihood of successful completion.
- There are inherent difficulties in a study with a large number of service providers. These include management of the study and homogeneity of intervention. The number of providers should be minimised where at all possible.
- A flexible approach to the chosen methodology is desirable so that the approach can be modified should circumstances require it.
- The inclusion of a qualitative element into a quantitative study will complement the quantitative findings and may help to save the study should serious difficulties arise. It will also provide additional information to contextualise the quantitative data, as for example, in the pilot study, see Section 5. The qualitative aspect can take the form of semi-structured interviews conducted at the time of collection of quantitative data and this has been included in the protocol set out in Section 10.

Acknowledgements

We would like to thank Dr Bente Berget and Professor Bjarne Braastad of the Norwegian University of Life Sciences, Ås; and Marjolein Ellings of Wageningen University, the Netherlands, for their help and cooperation.

5 Selection of Outcome Measures: Results of a pilot study

A pilot study of a series of outcome measures of health and well-being was conducted in order to:

- test the acceptability of those measures to study participants
- identify any difficulties associated with their use
- assess, through interviews, whether the dimensions of health and well-being addressed by the methods were considered to be appropriate ones
- explore the health and well-being of a sample of existing garden project clients using the outcome measures and interviews (as a 'snapshot')
- estimate the potential 'effect size' of the intervention of attending a garden project by examination of the results from the outcome measures and also by comparison with published values
- establish the statistical dispersion (Standard Deviation) of the results from the outcome measures to aid the calculation of sample size for a trial
- identify other areas or measures of health and well-being that would be suitable for inclusion in a future trial of therapeutic horticulture

It is important to note that it was not the aim of the pilot study to carry out a small-scale trial of effectiveness.

Approval for the pilot study was obtained from a NHS Research Ethics Committee (REC) and all conditions and protocols were followed as specified in the (amended) proposal approved by the committee.

Selection of Outcome Measures

Five potential outcome measures were selected for study after examination of the literature and consultation with psychiatrists and researchers in the field of

mental health. Outcome measures were chosen to assess participants' underlying symptoms and problems and also to explore those dimensions of health and well-being that may be influenced by attending a garden project. Suitable outcome measures need to be sensitive to the effects of the intervention. Our past research has shown that STH can promote social inclusion, provide social opportunities, develop skills in both social functioning and task-related activities and to improve the subjective sense of well-being (see Sempik *et al*, 2005). The selected outcome measures, described below, address aspects of health and well-being arising from those particular dimensions. For example, social functioning is explored by the WHOQOL BREF Domain 3 and by the CORE 'life functioning' dimension; development of skills and confidence in personal ability is explored through the general self efficacy measure.

The outcome measures selected were:

- **Clinical Outcomes in Routine Evaluation (CORE) Outcome Measure**

The CORE Outcome Measure is part of an assessment and quality evaluation system for psychological therapy services (CSG, 1998; see also: Connell *et al*, 2007). It addresses clients' 'global distress' across four dimensions (listed below). The measure has been trialled extensively and is sensitive to change. For example, Evans *et al* (1998) reported the effect of a short term counselling intervention on student CORE scores in their paper discussing reliable and clinically significant changes in measures of mental health outcomes. The four dimensions of CORE are:

- Well being (4 items).
- Problems or symptoms (12 items).
- Life functioning (12 items).
- Risk (to self or others) (6 items).

Although the authors of CORE originally considered these dimensions to be discrete and promoted the use of separate scores for them, in

addition to a total score (see CSG, 1998), recent work suggests that they are highly correlated and, in fact, measure the same construct from different perspectives (Connell, personal communication February 2007). Published data is available for normative and clinical UK samples (Connell *et al*, 2007).

- **The Hospital Anxiety and Depression Scale (HADS)**

HADS is a 14 item questionnaire which provides separate scores for anxiety and for depression (Zigmond and Snaith, 1983). Although originally intended as a screening tool for use in general out-patient clinics, it has been shown to be useful as a research tool. It is sensitive to both drug treatment and psychosocial interventions (see review by Herrmann, 1997). There is much published data on HADS, including normative data for the UK and the effects of gender and other demographic variables (Crawford *et al*, 2001). The depression subscale is based on the concept of 'anhedonia' (the inability to experience pleasure from events and activities usually associated with pleasure) and is a non-specific indicator of depressed mood and sensitive to mild disturbances (see Herrmann, 1997, p 33).

- **The World Health Organisation Quality of Life questionnaire – short form (WHOQOL – BREF)**

The WHOQOL BREF questionnaire is a shortened (26 item) version of the WHOQOL 100 item quality of life measure. It has four domains:

- Domain 1 – physical health incorporating facets of energy and fatigue, activities of daily living, mobility, pain and discomfort, work capacity and others (7 items).
- Domain 2 – psychological well-being, including bodily image and appearance, self-esteem, positive and negative feelings (6 items).
- Domain 3 – social relationships – personal relationships, social support and sexual activity (3 items).

- Domain 4 – environment including financial resources, participation in and opportunities for recreation and leisure activities (8 items).

The WHOQOL BREF has been validated by WHO centres around the world and data and scores are available for many different countries (see Skevington *et al*, 2004 for WHOQOL BREF data from 23 countries and over 11,000 participants).

- **General self-efficacy questionnaire**

'Self-efficacy' is a person's belief that they can complete a task successfully or overcome adversity (see Bandura, 1994). It is frequently applied to specific situations, for example, someone's belief that they can give up smoking (see, for example, DiClemente *et al* 1985). However, a generalised form of self-efficacy questionnaire has been developed by Schwarzer and Jerusalem (1995) as a 10 item form. The authors have also made available, for download from their website⁴, a raw data file (in SPSS format) of the results of 18,000 respondents from around the world. This data was used to calculate the mean self-efficacy score of adult respondents from the UK and compare it with the scores obtained from the pilot study. The format of the general self-efficacy questionnaire allows additional questions to be included which are specific to the context of the research. Such questions were not included in the pilot study but may be added to the outcome measure for use in a future trial (see later for discussion).

- **Use of services and substances**

A form was constructed to record the use of GP and hospital services; interaction with the criminal justice system; use of alcohol, cannabis and other substances. This form was based on one used by the Oxford Mental Health Initiative to record service and substance use by members of a therapeutic community (Rex Haig and Yollande Hadden,

⁴ <http://userpage.fu-berlin.de/~health/selfscal.htm>

personal communication). The form incorporated the 'FAST' alcohol screening questionnaire (Hodgson *et al*, 2002) for questions relating to alcohol use.

Four dimensions were covered by the form, these were:

- Use of medical and social services (GP, psychiatrist, hospital, social worker and time lost to illness – 7 items).
- Interaction with the Criminal Justice System (arrests and convictions – 3 items).
- Use of Alcohol (the 'FAST' questionnaire – 4 items).
- Use of cannabis and other substances (as named/specified by participants – 3 items).

Administration of outcome measures

The five outcome measures were ordered in one repeating Latin square format to prevent any effect due to the sequence of questionnaires.

Additionally, a form was used to collect demographic information – age, gender, frequency of attendance at a garden project, duration of attendance, whether in employment; and details of illness – diagnosis and duration.

A brief introduction to the questionnaires was contained on a cover page and the forms were presented as a single booklet. Participants were provided with a separate information sheet and consent form. The forms were placed by participants into an envelope and sealed to protect anonymity. The outcome measure forms did not collect participants' names. Consent forms were collected separately.

Participants

29 participants were recruited from four garden projects. Additionally, 24 subjects (employees of a different university to the researcher's) with no reported history of mental ill health were recruited to a 'control' group for comparison purposes.

Participants from the garden projects were interviewed after they had completed the questionnaires in order to collect information on any difficulties encountered with the forms; their perceptions of the relevance of questions asked and what additional information, in their view, needed to be collected.

Results

Interviews: ease of completion and relevance of outcome measures

In general, the questionnaires presented no difficulty to participants, although a small number required help from project staff. The time taken to complete all of the forms ranged from twenty to forty five minutes. The format and layout of the questions were well received, particularly, since many of the participants were familiar with this method of assessment:

“I found it quite straightforward to do. I have done these sort of forms before when I was having cognitive behavioural therapy. We used to have to regularly fill in these, similar type of forms, so I was quite used to these questions and the layout and the format.

...There were a few questions which I did mull over but came back probably to my original thoughts”

“I found it quite user friendly really. I went through it all quite quickly. It made a lot of sense, yes.”

“I like the ones where you just tick a box and you have, like, “not at all”, and it works it up from bad to best, or best to bad. I like those ones because I find those quite easy because it says it at the top of the page.”

The participants considered that the questions addressed issues of health that were important to them, for example:

“I found that form very helpful, because the questions on it are very, very based on how I feel at the moment”.

One issue that did arise was the time period referred to by the questionnaires (for example, the WHOQOL’s use of ‘in the last two weeks’ or ‘over the last week’ used by CORE) since many participants experienced pronounced swings in mood, for example:

“I think personally, the forms, er, asking for the relevant information, over, say like, a period of a week, I don’t think I could honestly realise – I don’t think it’s a long enough span of time to ask because things vary so greatly with

If you’ve had, relatively, not too bad a week, obviously, you answer then to that degree. A week down the line you could get it a hundred percent different. But that’s how questionnaires are.”

“Some of them, I guess were a bit, a bit difficult. Because I suffer from rapid cycling bipolar disorder so my mood can change rapidly within a day. So it can ask you a question but you can never know what to answer, whether you’re - depending what mood you’re in, does that make sense? You have two different answers, so that was a bit difficult. I have two different answers because I have two different moods.”

These comments also raise an issue regarding the difference in approach used by the CORE and HADS instruments and that used by WHOQOL BREF and the general self-efficacy questionnaire. The former ask for the incidence or occurrence of a mood or feeling within the time period, for example:

Question: I have felt terribly alone and isolated (CORE question 1)

Responses offered range from ‘not at all’ to ‘most or all of the time’

Whilst questions in the latter refer to mood or activity over the whole time period, for example:

Question: How well are you able to concentrate? (Q 7.)

Responses offered range from 'not at all' to 'extremely'

However, there were few missing values in the responses collected suggesting participants had been able to decide on responses in spite of the conflict caused by fluctuations in mood during the period referred to by the instruments. The following examples illustrate how two participants had overcome the difficulties in completing the questionnaires related to changing moods over time:

“It was OK. I tended to look back on, say, on the last week and look objectively and say that was my worst point, er, in that week and sort of answer it in that way because of my moods and everything fluctuate continuously or thoughts or whatever, behaviour, and so I tended to stick to one level and say well that’s how I was feeling at my worst and kept it at that all the way through. So, I hope that is what you were looking for.”

“The questions say were very specific, say, ‘over the last week’ and my particular illness is very, sort of, up one minute and then low the next. And different outside factors affect the way I feel and think, erm, but I know, for instance, I obviously tried to take a broad aspect or broad picture over the last week. But I wasn’t here last week, and because I was not well enough to be here, and my moods were very low, erm, but making myself over the weekend, erm, plan to do things to increase my mood, to lift it again, as I know the fact that today I’m coming to [the project] has, erm, sort of, given me the encouragement and the, erm, sort of ability to enable myself to get here knowing that as soon as I’m here those moods have completely lifted. I mean now, I feel perfectly fine, whereas last week I felt awful. So it’s, the questions were very, sort of, erm, cos as I say, I think I tried to take my feelings or thoughts when they were at their worst, so that’s the only thing that you, maybe, you

need to clarify a little bit more on the questions.”

None of the participants objected to, or felt uncomfortable with, any of the sensitive questions presented to them. A number of minor issues was raised or observed regarding the forms, for example, instructions provided as part of the forms were confusing to some - one respondent reported that he found the “for office use part of it a bit confusing” and was unable to complete it.

Fourteen of the 29 respondents in the garden group completed the two sample questions intended for illustration purposes on the WHOQOL form.

Suggestions for additional questions and changes

Many participants wanted to include a questionnaire or questions regarding participants’ satisfaction and experience with the garden project itself, as the following quotes illustrate:

“The questions on the health on the form are good, but maybe there should be a bit more about the projects that you attend and how you feel about attending them and that sort of thing.”

“It seems that it’s more to do with everyday life than directed to the project.”

“Perhaps, it’s about the gardening...Maybe, what you hope to get out of it, out of the project and perhaps future goals, I don’t know if you can do that, about the future.”

“Maybe it didn’t make enough reference to; or did it? Quote me if I’m wrong, about exactly , what does this project actually, how much does it actually mean to you...and how does it actually impact on your life. Why the person that’s answering the paper – it could help them, why it could help them? Why come here? To a place like this? Why not go to, say, some other manual place? I don’t know, but..so to come to a place

like this, if you don't like it you go. If you stay, you stay. Why do you stay? That wasn't asked."

What kind of questions would you ask?

"Just the basic questions, I suppose, about how you feel, about coming to [the project] and how it make you feel, and how you cope with that, things like that I suppose."

Two participants considered that questions on smoking habits should be included as part of the 'use of services' form as attending the project had reduced their smoking (the research literature suggests a complex link between smoking and mental health).

One suggested clarification was required regarding the use of the term 'drugs' and suggested that use of prescribed medication could also be of interest to the researchers:

"It goes onto what other drugs you have use. I got slightly confused on that. But then it's not the normal drugs [medication] that you would take. It's other drugs....it might also be interest for you to ask what medication we're on."

Perceived benefits of the projects

The wish to include questions about the project may derive from the participants' perception that it was the project itself that was responsible for improving their health and well-being or maintaining the fragile *status quo* of these. The questions, whilst addressing their mood, did not acknowledge the fact that, in their perception, the garden surroundings (and associated factors i.e. staff, social opportunities) were exerting a powerful effect on it. The following quotations illustrate this:

"That's why it's here. That's what it's all about, to improve your quality of life. It gives you some meaning in life. Something you can be proud of,

you know, that you're really achieving something in life, and some aim in life.

And without it. I would be a lot worse off. When this closed down in July, and none of us we're sure it would open again – it was devastating. Not just for myself, I can speak for all of the other volunteers. It had a massive impact. Believe you me.

I mean it really takes you. You might be, if you want to put it on a scale, you might be, ...you might be halfway up the ladder, something like that hits you – you're back down the ladder again, and thinking, starting to think negative thoughts again, about life in general, and you know.

It just gives you a real buzz, a real sense of 'Yeah, you've got something in life to work for, to look forward to. Your life's not done. Like, sometimes, I, you do tend to think like that. Sometimes I get bouts of I wish I wasn't here. Sometimes I think, I, you'd probably be better off if you wasn't. I don't mean here, I mean in life in general. But that's what I'm saying, you get these depressions and negative thoughts. But something like this can take that away from, it gives you something to focus on, your mind's focussing. I have a lot better days when I'm working here, a lot, lot better days than if I'm at home."

"It does a lot for me because I live on me own and I don't see me family much. They're out at work and they've their own lives. I'm on my own so, I come out here and I feel a lot better. Makes me feel well in meself. It's another world to me."

Demographic data

There was an even distribution of genders in both samples; and similar age profiles with a mean age of 40.5 for the control group and 45.2 for the garden project group.

The garden group attended the project, on average, twice weekly and the mean duration of attendance was 3.4 years, although the minimum was only four weeks and the maximum, 15 years. The mean duration of illness was 13.3 years (n= 26, SD = 10.5)

A self-reported clinical diagnosis was available for 21 participants and this is shown in Table 5.1 (below), Table 5.2 shows the self-reported diagnosis from the *Growing Together* study of social and therapeutic horticulture (Sempik *et al*, 2005) for comparison. These two sets of data illustrate the heterogeneous nature of the client group – a spread of conditions was observed in both studies. However, the prevalence of different conditions was not the same. For example, psychosis/schizophrenia were the most frequently reported conditions in the *Growing Together* sample (38.8%) but only accounted for 6.9% of clients in this pilot study. Similarly, 24.1% of participants in this study reported bipolar disorder compared with only 6.1% in the *Growing Together* sample. The incidence of depression was similar in both samples, with 24% reported in this pilot study and 20% in the *Growing Together* study.

Table 5.1: self-reported diagnosis – pilot study

Diagnosis	Number of participants	Percent
Anxiety with depression	3	10.3
Depression	7	24.1
Depression with anxiety	2	6.9
Bipolar disorder	7	24.1
Schizophrenia	2	6.9
Not specified	8	27.6
Total	29	100

Table 5.2: self-reported diagnosis – *Growing Together* Study (Sempik et al, 2005).

Diagnosis	Number of participants	Percent
Anxiety disorders	9	18.4
Bipolar disorder	3	6.1
Depression	10	20.4
Psychosis	3	6.1
Post traumatic stress disorder	2	4.1
Schizophrenia	16	32.7
No diagnosis	6	12.2
Total	49	100

The Outcome Measures

Using the five outcome measures described above, the garden group participants (who had been attending projects for a mean of 3.5 years) showed significantly poorer scores for mental health and well-being than the control group (and also poorer scores than published normative sample data where available). However, interviews with the participants, extracts of which are shown above, strongly suggest that attending the project had brought perceived benefit to them and may contribute to their strategy of coping with their condition. Indeed, when services had been withdrawn, this has caused them some considerable distress. Baseline data were not available for this pilot study, so it is not possible to know the magnitude of any change in outcome measures that would have been observed had it been possible to collect such data. However, this study, as intended, collected the statistical dispersion of the data i.e. the Standard Deviations of the scores. By using the values for the Standard Deviations of the scores and realistic, but conservative, assumptions of effect sizes (now that values for the actual state of health and well-being are known) sample size calculations could be performed. These are presented after an examination of the individual outcome measures.

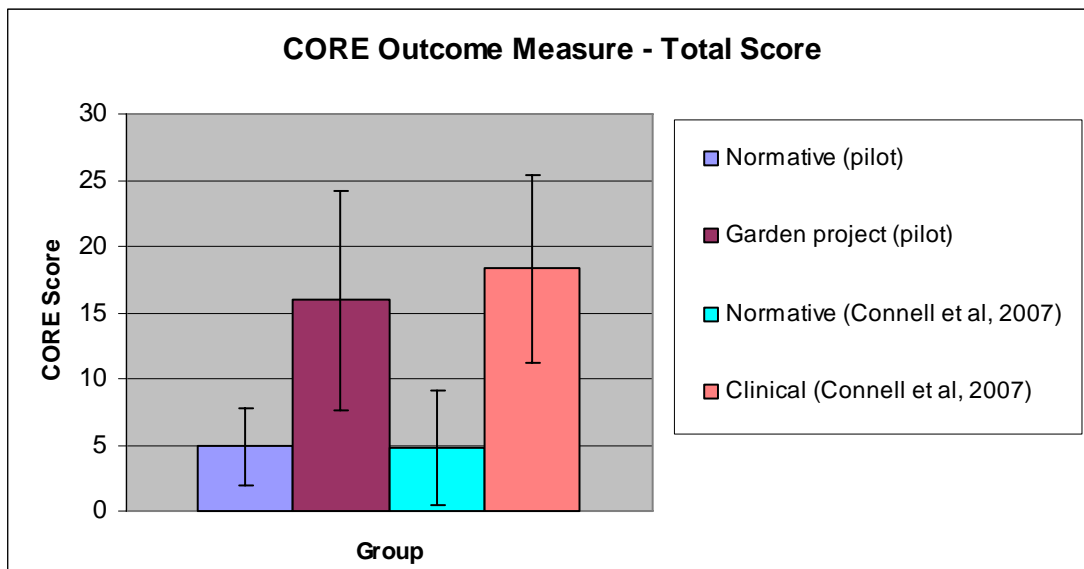
CORE

The results from the CORE Outcome measure for the garden group and the control group are shown in Table 5.3, below, and also for clarity in graphical format in Figure 5.1. Published data (Connell *et al*, 2007) for normative and clinical samples are also shown. The garden group had a significantly higher ($p=0.000$) mean score than the control group; and a higher score than the published value for a normative UK sample (indicating poorer mental health), but a slightly lower score than that for the published clinical sample. The control group had a similar mean score to the published normative value for the UK.

Table 5.3: CORE Outcome Measure responses from garden project clients and from a control sample; comparison with published normative and clinical values

	Control Group (Pilot Study, n=24))		Garden Group (Pilot Study, n=29)		Normative sample (Connell <i>et al</i> , 2007) (n =535)		Clinical sample (Connell <i>et al</i> , 2007) (n=10,761)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Mean Total score	4.9	2.9	15.9	8.3	4.8	4.3	18.3	7.1

Figure 5.1: CORE scores for control and garden groups; and a comparison with published data



WHOQOL BREF

WHOQOL BREF scores from the pilot study are shown in Table 5.4 (below); garden group scores were significantly ($p=0.000$) lower than those of the control group across all of the domains of the WHOQOL BREF.

Data from two published studies are included for comparative purposes (Table 5.5). The first study (Skevington *et al* 2004) provides data from the UK as part of a validation study of the WHOQOL instrument involving 11,830 participants from 23 countries. Means and standard deviations of the whole UK sample are given, together with mean values for participants who were 'well' compared with those classed as 'sick'. Forty seven percent of the entire sample of 11,830 participants were classified as 'sick' although the number in the UK sample is not given.

Mean pilot study control group scores were higher across all domains than the published values for the whole UK sample or for the mean well or sick subgroups.

Mean garden group scores were lower for all of domains than the published values for the whole UK sample or for the UK 'well' subgroup. Garden group scores for Domain 1 (Physical health) were slightly higher than those of the 'sick' subgroup, but lower for the other domains.

Values for Standard Deviation from the garden group were similar to the published values (even though the value of n was substantially smaller) but control group Standard Deviations were lower than those published by Skevington *et al* (2004).

The study of Diehr *et al* (2006) examined the correlation of the symptoms of depression with indices of the quality of life. The participants in that study were patients who had been diagnosed with clinical depression in the primary care setting. Those with psychosis and bipolar disorder were excluded. Table 5.5 shows the mean baseline WHOQOL BREF values at presentation and those recorded nine months after treatment. Differences (improvements) are

seen in all domains and these range in magnitude from 0.7 to 1.4. When these are expressed as z-scores (i.e. in terms of the Standard Deviation of the population or grand mean), values range from 0.26 to 0.46. The change, therefore, approaches approximately half of one Standard Deviation of the population mean.

Table 5.4: WHOQOL BREF scores – pilot study data

Transformed scores (4-20)	Control Group (pilot study) (n=24)		Garden Group (pilot study) (n=29)	
	Mean	SD	Mean	SD
Domain 1 (Physical health)	16.62	1.80	13.37	2.77
Domain 2 (Psychological)	16.08	1.92	12.17	3.62
Domain 3 (Social relationships)	16.33	2.13	11.67	3.73
Domain 4 (Environment)	16.02	1.90	12.917	3.06

Table 5.5: WHOQOL BREF scores – published data

Transformed scores (4-20)	Skevington et al (2004) UK sample (n=475)				Diehr et al (2006) (depressed patients) (n=982)	
	*Mean (all sample)	SD (all sample)	Mean 'well'	Mean 'sick'	Baseline	at nine months
Domain 1 (Physical health)	15.8	3.8	15.4	13.1	12.16	13.52
Domain 2 (Psychological)	14.7	3.4	14.8	13.7	11.41	12.6
Domain 3 (Social relationships)	14.2	3.5	14.8	14.0	11.95	13.02
Domain 4 (Environment)	14.1	2.3	14.1	13.8	12.26	12.92

*corrected for age and gender

HADS

Table 5.6 (below) shows the HADS scores for garden and control samples. Scores of garden group participants were significantly ($p=0.000$) higher than control scores for both anxiety and depression; and higher than those of a published UK normative sample. The pilot study control group had scores close to published values for the normative sample. The authors of the scale, Snaith and Zigmond (1994), had suggested that scores of 8-10 were indicative of mild cases of anxiety and depression and values of 11-15 of moderate; and those above 16 of severe cases. There has been discussion in the literature regarding appropriate cut off levels for clinical diagnosis (see Crawford *et al*, 2001; Herrman, 1997) although the cut-offs suggested by Snaith and Zigmond (1994) are generally used in practice. Using these values, the mean depression score of the garden group (7.83) was just below the threshold for mild depression and well below the threshold for moderate depression (11.0). The mean score of the garden group for anxiety (10.72) was within the range for mild anxiety but just below the cut-off for moderate cases.

Table 5.6: Anxiety and depression scores (HADS) – pilot study and published data

	Control group (pilot study) (n = 24)		Garden Group (pilot study) (n = 29)		(Crawford <i>et al</i> , 2001) (Normative UK) (n = 1792)	
	Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation
Anxiety score	5.5	3.2	10.7	4.9	6.1	3.8
Depression score	3.3	3.4	7.8	4.5	3.7	3.1

General self-efficacy

The mean general self-efficacy scores are shown in Table 5.7 below. The mean score of the pilot study control group was significantly higher ($p=0.002$) than that of the published value for a UK sample (Schwarzer, 2007); and also significantly higher ($p= 0.000$) than the garden group. The mean score of the garden group was significantly lower ($p= 0.045$) than that of the published UK sample.

The garden group, therefore, had a mean self-efficacy score lower than that of the control group from the pilot study and also below that of a published UK normative sample. However, it was greater than the score reported by Berget (2006) for a sample of psychiatric patients at baseline in a trial of animal-assisted therapy.

Table 5.7: General self-efficacy scores – pilot study and published data

Control group (pilot study) (n = 24)		Garden Group (pilot study) (n = 29)		Normative UK Sample (published) (n = 219)		*Clinical Sample (Norway) (published) (n = 69)	
Mean	SD	Mean	SD	Mean	SD	Mean	SD
32.1	2.6	27.0	6.3	29.2	5.3	24.1	5.6

*pooled data calculated from Berget (2006).

Use of services and substances

The garden group accessed medical and social services significantly more often ($p=0.000$) than the control group and had significantly ($p=0.000$) more interaction with the criminal justice system in the form of arrests and convictions (Table 5.8). However, there was no significant difference in the use of alcohol or substance misuse. Although the control group had a higher mean score for alcohol use than the garden group, this was below the value (3) associated with 'problem' consumption of alcohol. Information on alcohol use was obtained using the 'FAST' alcohol questionnaire incorporated into the

'Use of Services and Substances' form (see above). This has only four items and is a shortened version of the World Health Organisation's 'AUDIT' alcohol questionnaire (see Bradley *et al*, 1998). It is possible that this instrument lacks the required precision for a study such as this. The use of the AUDIT form may be preferable in a future trial.

None of the participants reported using any illegal substances.

Table 5.8 : Use of medical services, encounters with the criminal justice system and use of alcohol and substances

	Control group (pilot study) (n = 24)		Garden Group (pilot study) (n = 29)	
	Mean	Std. Deviation	Mean	Std. Deviation
Use of Medical Services	1.5	1.5	6.6	3.8
Interaction with Criminal Justice System	.0	.0	.48	1.1
Use of Alcohol	2.3	4.7	.83	1.7
Substance misuse	.0	.0	.0	.0

How close are the mean outcome measure scores of the garden group to the mean scores of normative and clinical samples?

As mentioned above, the mean scores of the outcome measures from the garden group showed the group to have poorer health and well-being than published normative samples. Although baseline data are not available, it is useful to consider whether there could have been movement of the outcome measures of the garden group towards the normative distribution (and away from the clinical one).

One of the criteria of clinical significance of a psychotherapeutic intervention (discussed below) is the potential to restore an appropriate outcome measure of health and well-being of an individual to the normative distribution. It is useful, therefore, to examine the mean scores of the outcome measures of the garden group to see how close they lie to the published means for normative samples. This is shown in Table 5.9 below. The differences are expressed in terms of Standard Deviations (of the normative data).

Thus for some of the outcome measures, for example, general self-efficacy, the difference is relatively small (0.41 SD) whilst for others, such as CORE, it is greater.

Another criterion of clinical significance is movement of the outcome measure variable away from the clinical distribution. Table 5.9 also shows the differences between the garden group scores and those of published clinical samples. Again there is a variation in the magnitude of the differences.

This variation may have two possible causes:

- differential changes in the different domains or dimensions of participants' health and well-being (or different sensitivities of the outcome measures to different aspects of health and well-being)
- there are different degrees of overlap of the distributions of clinical and normative populations for the different outcome

measures, i.e. the differences between means of normative and clinical populations are different for different outcome measures

- a combination of the above

Published differences in the means of outcome measures from normative and clinical populations are summarised in Table 5.10. The greatest difference is in the CORE outcome measure and the smallest in the WHOQOL domains.

Table 5.9: Differences in scores of outcome measures between the garden group and published normative and clinical values

Outcome measure and domain	Difference between garden group and published normative values (given as SD of normative group)	Difference between garden group and published clinical values (given as SD of clinical group)
CORE Total score	2.58	0.34
HADS - Depression score	1.35	^a 0.7
HADS - Anxiety score	1.22	^a 0.1
WHOQOL - Psychological	0.74	0.28
WHOQOL - Social relationships	0.72	^b -0.07
WHOQOL - Physical health	0.64	0.4
WHOQOL - Environment	0.51	0.26
General self-efficacy	0.41	0.52

^a The Standard Deviation of the garden group was used to calculate these values as no other appropriate data were available; the clinical cut-off level of 11.0 was taken as the clinical value.

^b The negative value indicates that the garden group score in this domain was lower (i.e. poorer well-being) than the published value.

Table 5.10: differences between published normative and clinical scores of outcome measures of mental health and well being

Outcome Measure	Normative Sample	SD	Clinical Sample	Difference between clinical and normative (as SD of normative group)
CORE Total (Connell <i>et al</i>) ^a	4.8	4.3	18.3	3.14
HADS Depression score ^b	3.68	3.07	11	2.38
HADS Anxiety score	6.14	3.76	11	1.29
WHOQOL – Psychological ^{cd}	14.7	3.4	11.41	0.97
WHOQOL - Social relationships ^c	14.2	3.5	11.95	0.64
WHOQOL - Physical health ^c	15.8	3.8	12.16	0.96
WHOQOL – Environment ^c	14.1	2.3	12.26	0.80
General Self Efficacy ^e	29.2	5.3	24.1	0.96

Notes:

^a Clinical and normative values taken from Connell et al (2007)

^b Normative values taken from Crawford et al (2001); a score of 11 was used as the value for a clinical population for both anxiety and depression, this is the generally accepted cut-off value between mild and moderate conditions.

^c Normative values taken from Skevington et al (2004); clinical values are baseline scores recorded by Diehr et al (2006) for depressed patients.

^e Normative values for UK calculated from the data of Schwarzer (2007); clinical values are pooled data taken from Berget (2006), n = 69, of a sample from Denmark.

If there were no difference in the sensitivities of the outcome measures to an intervention (i.e. the intervention affected all of the dimensions tested) then the greatest change would occur for the variable for which clinical and normative distributions were furthest apart. In the data shown above, one of the greatest differences between a published clinical sample and the garden group lies in the General Self Efficacy score, yet for this measure the distributions are fairly close with just under 1 SD between the means of normative and clinical populations.

To obtain an approximate indication of the relative differences between the outcome measures, the differences between scores of the garden group and published clinical and normative values (shown in Table 5.9) were divided by the differences in mean published scores of clinical and normative samples (shown in Table 5.10). These values are presented in Table 5.11, below.

Table 5.11: relative differences between outcome measures

Outcome Measure	Index of difference between garden group and published <i>clinical</i> values	Index of difference between garden group and published <i>normative</i> values	sum of indices
CORE Total (Connell <i>et al</i>)	0.11	0.82	0.93
HADS Depression score	0.29	0.57	0.86
HADS Anxiety score	0.08	0.95	1.02
WHOQOL – Psychological	0.29	0.76	1.05
WHOQOL - Social relationships	-0.11	1.13	1.02
WHOQOL - Physical health	0.42	0.67	1.09
WHOQOL – Environment	0.33	0.64	0.96
General Self Efficacy	0.54	0.43	0.97

The first column represents the standardised difference between the garden group and the clinical sample i.e. a movement away from a clinical population so a greater value represents a greater ‘effect’. The second column represents the standardised difference between the garden population and a

normative population. A smaller value, in this case, represents a greater 'effect'. The third column, which is the sum of the previous two, acts as a 'checksum'. Since any movement away from one population represents movement towards the other, the sum should therefore be '1'. This is approximately so for all the outcome measures.

The data suggest that the range of outcome measures used in this pilot study are likely to have different sensitivities to an intervention such as social and therapeutic horticulture and should represent a broad view of health and well-being. From the data it appears likely that the greatest effect may be on general self-efficacy. Indeed, Berget (2006) found that particular outcome measure to be the most sensitive in her study of animal-assisted therapy. **However, it is important to note that the arguments and observations presented above are based on the findings of many different researchers operating in a number of different countries and so caution must be exercised in the interpretation and use of the data.**

What are 'clinically significant changes' in outcome measures?

It is important to consider the question of 'clinically significant changes' in mental health. Although the purpose of this concept has been to identify clinically relevant changes in *individual* clients based on statistical data from a sample, it is useful to examine this concept when considering a future trial of STH. Although it may be possible to design a trial to detect very small changes in outcome measures as a result of participating in STH, if these are not truly indicative of clinical improvement then such data will have little impact. Hence, it is necessary to establish what is the magnitude of a clinically significant change in the outcome measures used in the pilot study and then to relate this to calculations of sample size.

A clinically significant change can be defined in terms of the following question:

How does the end state of the patient compare with the scores observed in socially and clinically meaningful comparison groups?
(Evans *et al*, 1998; see also Jacobson and Truax, 1991)

This has been operationalised using three criteria:

- A. a change in outcome variables of at least two Standard Deviations from baseline to post test.
- B. a change in the outcome variable to within two Standard Deviations of a relevant normative sample.
- C. a greater likelihood of the outcome variable being in the normative distribution after intervention than in the clinical distribution. The following formula is used to determine the point at which this occurs i.e. the point where the Standard Deviations of the normative and clinical distributions are equal:

$$\frac{(\text{mean}_{\text{clin}} \times \text{SD}_{\text{norm}}) + (\text{mean}_{\text{norm}} \times \text{SD}_{\text{clin}})}{\text{SD}_{\text{norm}} + \text{SD}_{\text{clin}}}$$

(reproduced from Evans *et al*, 1998)

We can examine the data from the outcome measures shown above with respect to the three criteria for clinically significant change. For example, using the published CORE scores, we can calculate the cut-off scores for each of the criteria i.e. the scores to which *individual* clients need to move for a clinically significant change. These values are also expressed as changes in client scores (in terms of SD of clinical group score) to achieve those cut off values. These range from 0.69 to 2.0 SD (Table 5.12). Criterion C is the one often used (see Evans *et al*, 1998), however, for this outcome measure, a change of 0.7 SD i.e. just over half of one Standard Deviation, could be

regarded as clinically significant. For other outcome measures, where there is a greater overlap between clinical and non-clinical distributions even smaller changes may be significant. For example, the difference between clinical and normative mean scores for general self-efficacy is just under 1.0 SD. The first two criteria would therefore not be usable and under Criterion C a change of just under half a Standard Deviation (0.47) would give a clinically significant change.

Table 5.12: Published CORE data and calculated values for cut off scores of clinical significance

Criterion		Cut off score for clinical significance	Change in clinical score to achieve clinical significance (SD)
A	change of 2 SD	4.1	2.0
B	change to within 2 SD of normative sample	13.4	0.69
C	change to a position midway between normative and clinical distributions	9.9	1.18

Note

Assuming a normative mean score of 4.8 (4.3 SD) and a mean clinical score of 18.3 (7.1 SD).

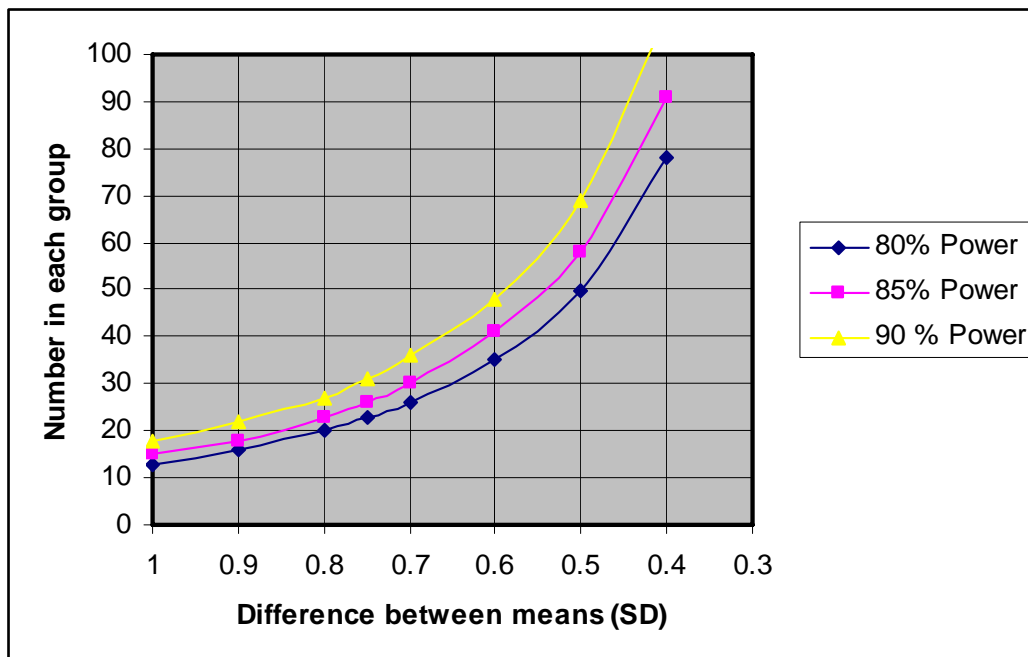
The observations presented above suggest that the response to STH is likely to be relatively small, possibly around 0.5 SD, although STH may cause larger effects on some outcome measures than others, for example, on general self-efficacy. A value of 0.5 SD is also around the level of clinical significance and so it would not appear to be fruitful to design a study to detect an effect smaller than this. It would therefore be prudent to design a study capable of detecting a change of around 0.5 Standard Deviations. We need to calculate the sample size required for such an effect size.

Sample size determination

Sample size calculations were performed using the computer program published by Lenth (2006). An alpha value of 0.05 was used. This is the probability of committing a Type I error i.e. appearing to detect a change when one does not exist. Calculations were carried out for powers of 80, 85 and 90% as indicated. These refer to the probabilities of detecting a true difference where one exists.

Figure 5.1 shows the relationship between sample size (groups of equal size) and effect size (expressed in terms of the Standard Deviation) for effect sizes ranging from 0.3 Standard Deviations to 1.0 SD. In order to detect an effect of 0.5 Standard Deviations, at a power of 80%, 50 participants are required for each of the intervention and control groups. A further 30 participants would be required for an effect size of 0.4 Standard Deviations.

Figure 5.1: the effect of the difference in means (expressed in terms of the Standard Deviation) on sample size (groups of equal size).



Realistically, 0.5 SD can be taken as the limit for a study. Indeed, Berget (2006) observed a statistically significant change of around 0.5 SD in general self efficacy score six months following a period of animal-assisted therapy. Such an effect size would require a sample, for each of the intervention and control groups, of 50-70 participants, depending on the power of the study. Decreasing the effect size to below 0.5 SD would increase the required sample sizes to unattainable values. Additionally, as discussed above, the clinical significance of an effect size of less than 0.5 SD of the outcome measures would appear to be questionable. **Hence, a study protocol that aims for an effect size of 0.5 SD of the outcome measures described in this section appears to be the most appropriate.**

Summary

- Garden project clients represent a heterogeneous group with regards to diagnosis of mental ill health.
- The instruments used in the plot study are acceptable to clients. They appear, in the view of participants, to probe appropriate areas of health and well-being.
- Clients would like to see additional questions relating to their experience of the project and its significance to them (these could be included within the self-efficacy instrument).
- Some clients suggested the inclusion of questionnaires of smoking dependence, an instrument such as the Fagerstrom test of nicotine dependence) could be included in the Use of Services form.
- The precision of the test of alcohol use could be improved by replacing the FAST screening test with the AUDIT test.
- Clients attending a garden project have poorer mental health and well-being than a corresponding age and gender matched control and than published values in the literature. This is the case even for clients who have attended a garden project for a relatively long period of time (mean 3.5 years).
- The health and well being of clients at entry into a garden project (baseline) are unknown.
- Clients reported that they perceived that attending projects was beneficial to their mental state of health and psychological well-being. Withdrawal of the service caused them distress.
- The effect size (of attending a project), as measured by a range of outcome measures, may be small. A protocol designed to detect an effect size of around 0.5 Standard Deviations is likely to be appropriate.
- A sample size of 50 participants per group would be required for a study of 80% power at an alpha value of 0.5 when the effect size is 0.5 SD.
- Effect sizes of less than 0.5 SD would require unrealistically large samples and such changes in the outcome measures may not have any clinical importance.

6 Assembling the sample

In the previous section we have considered the sample size necessary for a trial of STH and concluded that an appropriate number would be 50 – 70 participants per group (i.e. intervention and control), depending on the chosen power of the test. But would it be possible to assemble such a sample from garden project users?

One of the main difficulties of other researchers attempting controlled or quantitative longitudinal studies of therapeutic horticulture (or related areas) has been recruitment of participants (see Section 4). This has led to studies being carried out with reduced statistical power or to the change in methodological approaches once a study has started. Because of the constraints (financial and other) on many researchers, it is often necessary for them to start assembling their samples, i.e. finding groups willing to take part in the research or agencies willing to refer participants to interventions only after funding has been secured and therefore once the research has actually started. Because of the fixed-term nature of most research projects, any difficulties in recruitment have a serious impact on the quality of the data as it is often impossible to extend the duration of the study. It is therefore important that recruitment mechanisms and commitments to participation are securely in place prior to the commencement of a main study. A survey of garden projects was therefore conducted to gauge the level of interest in participation in a future trial of STH and to obtain an initial commitment to a study.

Garden projects from across the UK that offered a service primarily for people with mental health problems were identified from Thrive's database and were sent a survey form and reply-paid envelope for their response. A reminder was sent by post four weeks later and a further reminder was sent by email to those projects for which an email address was available. The survey questionnaire asked whether the projects were willing and able to take part in research; how many new clients they received each month; how many places were likely to be available for use in a trial; how many days of attendance

were available for clients; their level of contact with referring agencies and estimated costs (to the project) of participation in a trial of STH.

The results of the survey are summarised in Table 6.1 below. One hundred and four projects were identified and contacted; and responses were received from 60 of them. Fourteen projects were visited during the course of the study to observe activities and to interview staff.

Table 6.1: Summary of survey of garden projects for inclusion in a trial of STH

Project Status	Number of projects	Percent
Identified and contacted	104	100
Responding to the survey form (or by email)	60	58
Willing to participate	24	23
Responding, but unable to participate	36	35
No longer active	12	12
At risk of closing or currently winding down	6	6
Active but unable to take part because of stated reason	8	8
Suitable for inclusion	11	11

Twenty four respondents replied that they were interested in participating in a forthcoming trial of social and therapeutic horticulture. The projects were based both in the voluntary sector (16 projects) and NHS (8 projects). Some were not fully operational, so could not participate in the near future, while others were considered unsuitable, for example, those which offer formal psychotherapy and so operate outside of the paradigm of 'Social and Therapeutic Horticulture' which is discussed in Section 7. Those projects that

could only offer one day per week were also not included. In all, 11 projects were considered suitable for inclusion in a trial. Their details are summarised in Table 6.2. A total of 95 placements were reported as being available. It is possible, however, that this could be an overestimate of the true number of places. Also, there is likely to be a ‘wastage’ of placements due to dropout from the study. An attrition rate of around 25% can be expected (see Section 4). Therefore, two cohorts of participants would be required in order to achieve the required number of 100 participants. As each placement would be used for twelve months during the course of a study, the length of the intervention part of the study would be twenty four months, plus the time taken to recruit the first participants. A timetable and proposed protocol for a trial is discussed in Section 11.

Table 6.2: Garden project places available for a trial of STH

Ref	Sector	Affiliation	Number of Garden Places	Number of Days for Each Place
37	Voluntary	Associated with larger UK charity	10	3
58	Voluntary		2	2
59	NHS		4	2
69	Voluntary	Independent charity	4	1-2
81	Voluntary	Independent charity	5+	3
84	Voluntary	Associated with larger UK charity	10	1-2
96	Voluntary	Associated with larger UK charity	20	2
97	NHS		3-5	2+
98	Voluntary	Independent charity	10+	1-2
103	Voluntary	Associated with larger UK charity	20	3-5
104	Voluntary	Independent charity	5+	3+

Barriers to participation in research

As described in Section 2, most staff at garden projects believe that there is a need for more research into STH and support such work. The relatively small number that were willing to participate in research, however, does not indicate that the others did not value such work on their behalf. A number of projects had had difficulties with funding or referrals. For example, 12 were no longer active and six were in danger of closing. This is illustrated in the following quote from the manager of one project:

“Sorry I did not reply to your last correspondence, we are under review and at present have an uncertain future. As much as I would love to be a part of your research I don't feel we can commit to anything as I don't want to get half way through and let you down mucking up your outcomes. However I would still be interested to see more information on what you plan to do and the type of research you are doing. Who knows if we get this sorted we may be able to work together in the future.”

The most common reason given by an active project for being unable to participate in the research was a shortage of staff, for example:

All staff are part time and too busy to take on any more work”.

“Whilst I would love to help by participating in your survey I am afraid that our already overstretched resources do not allow us to assist you at this time”.

Additionally, confidentiality of data was an issue for one project manager:

“At present we have restricted staffing and under our present operational policies, we are presently restricted from providing access to service users/information to third parties. [***] healthcare NHS Trust only provides access to services directly linked to our service user providers”.

It is possible that such considerations may also have played a part in the lack of response from those projects that did not return their survey forms. It is not clear how many of those are still active and how many have closed.

Advisory Group – a ‘community of interested parties’

Representatives (including service users) from all of the projects who expressed an interest in participating in the research (including those that were not included in the final list), and individuals with a specific interest in research in the field of STH (for example, other researchers, occupational therapists and practitioners) were invited to be part of an Advisory Group for a future study. Two meetings of the group were held during the course of this twelve month study and a useful dialogue was established. Other projects and researchers were consulted through the establishment of an informal ‘community of interested parties’ which includes projects that are willing to participate in a study and other researchers and practitioners with an interest in mental health and nature-based approaches who are able to offer useful advice. The community operates through informal meetings and visits and so contact between the researcher (JS) and other researchers, practitioners and STH project workers is maintained. The creation of such a community has been a time consuming process, but it will play an important role in the eventual process of recruitment of participants to a study. By establishing such a community prior to starting a full scale trial of an intervention such as STH, many of the potential difficulties of a trial can be identified at an early stage and solutions can be devised. As mentioned in Section 4, carrying out a preparatory exercise before conducting a fixed term study has advantages.

Summary and Recommendations

- Twenty four garden projects expressed an interest in taking part in a study of STH and 11 have been identified as 'suitable' for inclusion in such a trial.
- Ninety five placements were reported as being available.
- Taking into account any overestimation of places and attrition, then two cohorts of participants (for an intervention duration of 12 months) would be required in a study. The intervention phase of the trial would therefore last for 24 months (plus any initial recruitment or waiting time).
- An Advisory Group and 'Community of Interested Parties' has been established in preparation for a full scale study.

7 Inclusion and Exclusion Criteria

Clients with mental ill health who attend STH garden projects have a variety of diagnoses (as was seen in the pilot study for the selection of outcome measures, Section 5). Also, there is no evidence to suggest that clients with any particular condition respond better to STH than others. Therefore, in an initial trial of STH there is no reason to exclude participants on the basis of diagnosis alone. The issue of including participants with a range of diagnoses in a study was discussed with a group of STH project managers. The unanimous view among the managers was for the greatest possible inclusion, since a heterogeneous sample would reflect the heterogeneous nature of clients attending a project. This, in their view, would make the results of a study more generalisable to 'real life' and more acceptable and meaningful to policy makers. The study is aimed at adults, since children have different needs and this is reflected in a different approach in the use of STH and horticultural therapy for children. Whilst there are a number of projects that specifically provide places for children with learning difficulties and developmental disorders such as autism, there are very few places for children with mental ill health (as distinct from disaffected children or those excluded from school). None of the projects that registered an interest in taking part in a trial of STH provided a service for children.

Suggested inclusion and exclusion criteria are summarised in Table 7.1 below.

Table 7.1: Suggested inclusion and exclusion criteria for a study of STH

Subject	Inclusion	Exclusion
Participants	Persons with mental ill health who have been referred to a therapeutic garden project (or who have self referred) and who would be offered a place at a day centre or STH garden project under usual circumstances.	Persons with learning disability alone; other disabilities or conditions which would prevent participants completing outcome measure questionnaires.
Age	Adults, aged 18 or over	Children and young people
Ability	Able to complete outcome measure questionnaires (including with assistance).	Lack of sufficient awareness/comprehension to complete questionnaire even with assistance.
Diagnosis	Clients attending garden projects form a heterogeneous group with respect to diagnosis. A heterogeneous sample, therefore, would be appropriate to match the situation found in under standard working practices at STH garden projects; this would include persons with dual diagnosis (substance misuse and mental ill health); persons with personality disorder; persons with mental ill health and learning disabilities.	There will be no exclusion on the grounds of diagnosis alone.
Involvement with a garden project	Persons who are not currently involved with a garden project, gardening group or who list horticulture or gardening as a major hobby or leisure activity.	Persons who are currently involved with a garden project, gardening group, community garden or who list horticulture or gardening as a major hobby or leisure activity; persons who have attended or been involved with a garden project in the past.

8 Defining the Intervention and Control Activities

The practice of social and therapeutic horticulture (STH) is a varied one. It takes place in a number of different types of location (for example, gardens, allotments, farms and outreach i.e. offsite work); and with clients who have many different needs. The service is delivered by practitioners who have a variety of different skills, training, backgrounds and qualifications. Although described as *horticulture*, many different activities are also present. These range from tasks usually associated with gardens, such as landscaping and construction of sheds and garden buildings, through to arts – painting, drawing and ceramics; and crafts such as blacksmithing and weaving. Yet, these clusters of apparently disparate activities retain a coherence that is centred on the garden and the identity of the project. A visitor to a therapeutic garden will recognise the characteristics and *nature* of what he or she sees, but will struggle to define them. After a number of visits to different gardens they will identify common themes or approaches; they will have an impression of what (*for them*) constitutes ‘social and therapeutic horticulture’, yet will have difficulty describing that paradigm. In order to conduct a controlled trial of STH it is, however, necessary to define or standardise the *intervention*.

Standardisation would imply a prescriptive approach to practitioners delivering a service, yet the character of therapeutic garden projects is a fluid one. Activities, plants and even fellow participants and staff are often selected by a democratic process. Indeed, *political engagement*, as a dimension of social inclusion and implying a sense of control by clients over their environment, was one of the processes associated with garden projects that was identified by the *Growing Together* study (Sempik *et al*, 2005). If it is impractical to standardise the intervention, then it is necessary to define it sufficiently for use in a trial (and by other researchers) and also sufficiently for it to be reproduced widely should a trial show it to be effective.

Social and Therapeutic Horticulture (STH) is distinct from *horticultural therapy*. In our literature review (Sempik *et al*, 2003) we wrote:

““Horticultural therapy is the use of plants by a trained professional as a medium through which certain clinically defined goals may be met”.

“Therapeutic horticulture is the process by which individuals may develop well-being using plants and horticulture. This is achieved by active or passive involvement” (Growth Point, 1999, p. 4).

The distinction being that horticultural therapy has a pre-defined clinical goal similar to that found in occupational therapy whilst therapeutic horticulture is directed towards improving the well-being of the individual in a more generalised way. This can be the attainment of employment, an increased sense of self esteem or some other perceived benefit. Perhaps the term ‘social and therapeutic horticulture’ best describes the process by which plants and horticulture are used to develop well-being since, as this review will show, *social* interactions and outcomes play a significant role.” (Sempik *et al*, 2003, p 3)

These short definitions of therapeutic horticulture and horticultural therapy were produced by practitioners after many hours of discussion and debate. Yet they are still vague, perhaps deliberately so to allow practitioners and clients the room to manoeuvre in order to create an intervention to suit their individual needs.

However, we recently defined STH in the following way:

“STH can be seen as the participation by a range of vulnerable people in groups and communities whose activities are centred around horticulture and gardening. STH is distinct from domestic gardening because it operates in an organised and formalised environment.”
(Sempik and Spurgeon, 2006)

This allows us to narrow the focus and be more specific. It is important to state that the reason for excluding any particular approaches or projects from

a study is to maintain a homogeneity within the intervention and not because those other approaches are not effective, useful, or valuable.

Observing therapeutic ‘garden’ projects

Forty projects in the UK were visited during the course of this research and the *Growing Together* study. These were a cross-section taken from Thrive’s network database and included allotments, gardens, nurseries, daycentre garden projects, city farms, ‘educational/training’ farms and care farms. Additionally, 2 care farms were visited in the Netherlands and 2 in Norway. Over 2,000 photographs were taken. These projects represent the broad paradigms of horticultural therapy, therapeutic horticulture, social and therapeutic horticulture (STH) which is the intended intervention for our research and care farming. A number of dimensions were identified. These can be used to exclude those projects which lie outside the research paradigm of *Social and Therapeutic Horticulture*. ***It is important to stress again that the exclusion of projects is for the purpose of establishing as homogeneous an intervention as possible for use in research and not because those projects are not effective or useful.***

The following dimensions were identified:

Therapeutic intent and practice – therapeutic garden projects are *intended* to promote mental and physical health and well-being in their clients who may have mental, physical or social problems. There is an accepted and organised practice of social and therapeutic horticulture which is recognised by health professionals, researchers and others. There are training standards and qualifications for individuals, in addition to quality standards for gardens. This distinguishes STH from other forms of gardening, such as domestic or amenity gardening or even community gardening. Whilst these latter forms undoubtedly provide many people with opportunities for improving their health and well being; and may offer social opportunities, they are not *social and therapeutic horticulture*.

Location – the activities take place in a garden, allotment or other physical location, this distinguishes the intervention from ward-based occupational therapy; and also from outreach work and nature conservation. The presence of a ‘home’ location enables clients to form a bond with a specific location and develop a sense of place.

The natural environment – whilst horticultural therapy focuses on a person’s interaction with plants (which can be carried out in an enclosed room such as a hospital ward or therapy room), social and therapeutic horticulture takes a wider view of that person, **working with nature** (i.e. the horticultural dimension) and others (clients, staff and even visitors, i.e. the *social* dimension) in a natural setting such as a garden or allotment. Hence, the natural setting is a prerequisite for an STH project.

Democracy and involvement – we have previously identified that STH projects promote social inclusion through the dimensions of *production, consumption, social interaction* and *political engagement* (see Sempik, *et al* 2005). The dimension of political engagement is enacted through the clients’ opportunities to influence the activities of a project (suggest new ventures, for example). This degree of democracy is part of the intervention and this distinguishes it from similar activities *provided* as a service over which the client has no control. It is interesting to note that democratic decision making is also a defining feature of the therapeutic community.

Social coherence and community – STH projects foster the development of a community that works together, and socialises within the boundaries of the project (and occasionally outside). Staff and clients usually eat meals together and frequently prepare food together from produce they have grown.

Production – is an essential part of STH (i.e. the *production* dimension of social inclusion referred to in Section 1). It enables clients to develop a sense of identity as *workers* or *gardeners* and leave behind the label of *patient*. However, there is a range of emphasis on *productivity*. The theme of productivity but with ‘*lack of pressure*’ was an important one identified in our previous research so those projects which are focussed primarily on productivity and commercial activity lie outside of the paradigm of STH. So do care farms, which are generally intended to be productive. Indeed, many care farms started out as purely productive agricultural units before adopting a multifunctional approach to include the care and support of vulnerable people. Training specifically for employment can also be regarded as a form of production, the outputs being qualifications or paid work. So projects whose primary aim is training are part of a different paradigm.

Routine – the development of a routine and structure to the day is an important aspect of recovery for many people with mental ill health. The activities and procedures at STH projects are designed to facilitate such a development and therefore there is an expectation of commitment by the client to a regular, rather than a casual attendance. Therefore, those projects offering casual or ‘drop in’ activities should not be included.

Attendance: frequency and duration – our research has shown that most clients attend garden projects for two days per week or more frequently (see Sempik *et al*, 2005). It is possible that such a degree of attendance is necessary in order to achieve the benefits of STH, hence projects which are not able to offer at least two days attendance should not be included.

Although some garden projects provide a programme of fixed duration, most accept clients on an ‘open-ended’ basis; and most clients join projects with a long term view. In our previous study (Sempik *et al*, 2005) we found the mean duration of attendance to be 3.4 years

(n=121) and in this research the mean was 3.5 years (see Section 5). There is no information regarding the length of attendance required to produce benefit. Opinion also appears to vary among practitioners; some suggest that benefits are apparent after a few months whilst others recommend that at least a year is required for an effect. There is also likely to be variation between clients, depending on their condition or illness at entry into a project.

Many therapeutic community programmes run for 12 months. During this time the client members of the community will experience the full range of anniversaries, public holidays, birthdays and so on. With a similar period at a therapeutic garden, a client will additionally experience the change of seasons and engage in different aspects of work to coincide with those changes. Therefore, an attendance period of 12 months has been selected for a trial intervention.

Psychotherapy – STH projects have much in common with therapeutic communities (TCs) i.e. they are groups of vulnerable people who work as a community and run that community in a democratic way. However, STH projects do not provide group psychotherapy or psychoanalysis in the manner of TCs. Essentially, the view is that the way of life of the community, the social interactions, the (natural) environment and activities are the therapeutic dimensions. Therefore, those projects for which psychotherapy is an integral part should not be included.

Arts and crafts – various arts and crafts are a feature of garden projects and are generally associated with activities in the garden. Articles may be produced specifically for the garden. These may be practical or decorative, or contain both elements, for example, benches, wrought iron gates, statues and other items. Elements of the garden may be taken to inspire art or used as raw materials to produce it. Since such processes are an integral feature of many garden

projects, participants should not be excluded from them. However, the degree of engagement should be recorded.

In order to illustrate these different paradigms, a CD which contains a slide show⁵ of the different aspects of garden projects is enclosed with this report. The opening images show horticultural therapy/therapeutic horticulture in a hospital/occupational therapy context; the next section shows the STH environment and activities – the settings, opportunities for social interaction (and for working alone or independently within a social setting), horticultural activities and also art and crafts which utilise images and materials drawn from the natural environment; the final section shows the transition to an emphasis on productivity. This is a defining feature of care farms and is also present in some therapeutic horticulture projects which are engaged in larger scale commercial activities and exist as social firms or social enterprises. Many projects are involved in commercial activity in some form, for example, sale of plants and produce, but in general this occurs at a relatively low level.

Although these images are taken from many different projects, this visual essay illustrates the similarities that are observed. A small red dot in the corner of the photographs is used to indicate those images representing activities and settings lying within our chosen paradigm.

⁵ Double click on the file 'Defining STH.ppt' to automatically start the PowerPoint slideshow.

Control Activities

A control group is a comparison group against which an intervention is measured (see Section 3). This may be inactive, for example, a placebo pill in a drug trial, or 'treatment as usual' in the case of social care i.e. the control group receives no additional services or activities. It may also be an active substance or intervention. For example, Berget (2006) used 'treatment as usual' as the control for her study of animal assisted therapy whilst in an ongoing study of the effectiveness of care farms Ellings (see Section 4) chose to use work in bicycle repair shops and other small workshops as the activities for her control group.

In our survey of garden projects for this study we identified six projects (four within those considered suitable for inclusion as 'garden' projects and two in the wider group) that provided activities in a context that could be used as a control. A total of up to 32 places were reported as available. The following activities were identified:

- Information technology (IT): both training in IT and 'creative' use of the technology – use of the internet and computer applications including spreadsheets, drawing and image editing programs (e.g. Photoshop) and word processing.
- Woodwork: construction of wooden articles, frequently with a 'garden' theme i.e. garden tables and benches, nest boxes, decorative plant pot holders. These items were generally offered for sale in association with plant sales and so there was an element of repetitive production.
- Craft: pottery and ceramics; these activities were similar to those described above as part of the garden paradigm, however, they were used as the main focus for clients rather than being of peripheral interest.

These activities were provided in the context of the following dimensions described above: *social coherence*, *productivity*, *routine*, *attendance* and lack of formal *psychotherapy*. Clients engaged in them as their main activity or interest i.e. they chose whether to join a woodwork , computer or craft group.

However, where these activities were provided by a garden project, clients also had access to the garden.

‘Treatment as Usual’ vs. an ‘Active Control’

There are some difficulties associated with both *treatment as usual* and *active controls* in research on STH. Generally, there has been no attempt to define ‘treatment as usual’. It is assumed that study participants do not undertake any additional activities which may affect the outcome measures. Or if they do, then this is part of the process of appropriate treatment of the patients by their doctors and therapists (or of ordinary life) which is available to all – both control and intervention groups. Therefore, the issue is that ‘treatment as usual’ remains undefined.

Activities such as light assembly or repair work, computing and IT and woodwork are frequently offered to people with mental ill health because they are considered to be beneficial i.e. they provide occupation, distraction and recreation. There is an underlying assumption that this is an active intervention when used as a control, but there is no evidence or data to support such a claim and the magnitude of any effect is unknown. Therefore, if such activities are used as a control and both intervention and control groups show similar changes in outcome measure, the uncertainty will remain as to whether both changes were due to activity; or whether both changes were due to chance i.e. to a phenomenon such as regression to the mean.

Additionally, where control activities would be provided by garden projects there is an issue of ‘*contamination*’ i.e. participants in the control group have access to the same *natural* setting experienced by those clients engaged in horticultural activities. Although the amount of time spent in the garden setting is clearly different between the groups, it is still possible that this may have an adverse effect on the study and procedures to exclude control participants from the natural environment are likely to prove both problematic and unethical. Whilst it may be possible to identify specific activities for people with mental ill health entirely outside the area of STH, e.g. within the area of

sheltered employment, our attempts to recruit charities and organisations which provide such activities were unsuccessful. Continued perseverance could have yielded some results, but for a successful study it is important that those agencies taking part are wholly committed to the research, rather than being present on sufferance. If commitment relies solely on the 'lukewarm' promise of a small number of individuals who have been coerced into joining a study, then there is a danger that that the study will collapse (or that part of it will collapse) should there be any changes in circumstances or personnel.

There are also recruitment and ethical issues associated with a 'treatment as usual' control group. Whilst study participants may be indifferent to being in a control or treatment group for an intervention they are not familiar with or have not selected, they may feel quite differently if the activity is one they would actually like to do. This may have a negative impact on recruitment for a study. Additionally, many researchers consider it unethical to deny participants access to an activity or intervention which they perceive to have likely benefit. A solution to this problem is to offer the activity or intervention to all study participants but to place control subjects on a waiting list. The controlled study lasts for the duration of the waiting list. Setting the duration of the waiting list (and hence the intervention period) may present a challenge. It will need to be long enough for the intervention to exert an effect but not so long as to reduce recruitment or promote attrition. Berget (2006) used a 12 week intervention period and a 'treatment as usual' control group in her study of animal assisted therapy and found that there was no drop out after random allocation. Even without a waiting list, and the eventual promise of the intervention, study participants assigned to the control group were willing to be involved with the research for at least 12 weeks.

On balance, and in the absence of any data defining the control activities described above (and the effects of those activities on the selected outcome measures) it would appear to be safer to use a 'treatment as usual' control in preference to an 'active' one. A waiting list control would address any ethical issues that project staff may have with regards to denying services to their clients. Although a waiting list of long duration might benefit the researchers

by enabling a long term study, it is also likely to increase attrition. For example, it would be unreasonable and unethical to expect anyone to spend 12 months on a waiting list. Indeed, such a period would not be acceptable to garden project staff, however, a 12 week waiting list would be acceptable to most. Some projects already have waiting lists of 1-2 months for new clients and these could be used as part of the waiting list control period. Since Berget (2006) also used a 12 week timescale for her trial, using a similar duration will provide comparative data.

In defining STH (see above), we have suggested an intervention period of 12 months, but, as set out in the preceding paragraphs, there are limitations on the length of a control period. A trial of STH will, therefore, be comprised of a controlled part and a longer, pre-intervention/post-intervention longitudinal portion. Such a study design is presented in Section 10.

Summary and recommendations

- A paradigm of STH has been described which would be suitable as an intervention for a study.
- A waiting list control with 'treatment as usual' would be appropriate for a study of STH.
- A waiting list of 12 weeks duration should provide sufficient time for the intervention, whilst not being so long as to result in unacceptable attrition.
- A study protocol should contain a controlled part and a longer longitudinal portion (see Section 10).

9 Randomisation

In Section 3 we discussed the purpose of randomisation as a process for reducing bias in a study. However, not all studies are amenable to randomisation because of practical or ethical considerations. We have therefore explored the issues relating to randomisation in the context of a study of social and therapeutic horticulture (STH) for people with mental health problems.

Randomisation can be carried out at two different levels within the process of entry into a garden project:

- randomisation by the referring agencies (mostly community mental health trusts, CMHTs) of new clients entering their care. Participants, who give consent, can be allocated at random to either a garden project or to a control activity, for example, 'treatment as usual' or active control such as attendance at a daycentre or sheltered employment at a different location.
- randomisation of clients at the point of entry into a garden project to different activities available within the project, to a different schedule of activities or to a waiting list control (see Section 8).

The advantage of the first approach is that it removes or reduces the selection bias arising from clients' preference for a garden project (see Section 3) and conclusions from the study can be generalised to all those presenting to the CMHTs (and other agencies) with similar conditions. The difficulties associated with this approach is that it requires the referring agencies to carry out the random allocation.

The advantage of the second approach is that randomisation is carried out by project staff who have already shown an interest in the trial and have stated their intention to cooperate. However, there is a selection bias – only clients who show an interest and willingness to join a garden project are included in

the trial. But with regards to external validity i.e. the generalisability of the trial to the 'real world', it must be remembered that the actual willingness of clients *to attend* a garden project is an important prerequisite for effectiveness. This may be regarded as one of the 'indications' for the intervention. Whilst it is possible that clients who have little interest in gardens and have been randomly allocated to a garden project will show some level of compliance, those who have expressed a desire to join a project will probably show a great deal more.

Additionally, the presence of multiple sources of referral of clients to garden projects is an issue that needs to be considered. Four of the 11 projects shown in Table 6.2 (Section 6) received referrals only through their CMHT, however, seven, however, received referrals from multiple sources, including CMHTs, GPs, Housing Support Services, charitable organisation and self referral (five projects). For example, one project received referrals from eight separate agencies or organisations.

For random allocation at the first level (i.e. at entry into the CMHT process) to be successful and to proceed at a reasonable rate, it is necessary to have the cooperation of all (or most) of the referring agencies for each project. However, contact with CMHTs that referred clients to garden projects proved difficult and only one was willing to commit itself to participation in a trial of STH. This point of randomisation would also exclude self-referring clients. Self referral is an important method of entry into STH and some projects actively encourage this form of recruitment. For example:

“We encourage self referrals and like to think of our people as being introduced to the garden by other health care professionals or whoever else although we do ask for a reference from a health care professional. We get a lot of word of mouth intros as we target individuals who might use our service in our publicity rather than the professional body although we do make it available to them too” (STH project manager).

It is, therefore, unlikely that randomisation at the first level would be successful and produce sufficient participants for a trial. Whilst randomisation at that stage would enable generalisation to the wider population, random allocation at the point of entry into a garden project would not prevent the study from being a 'fair test', since in the 'real world' this may well be the sub-population for whom it is an effective intervention. The concept of 'personalised medicine' i.e. the production of remedies that are appropriate for specific sub-populations suffering from a condition or illness, rather than all those with the condition, is a rapidly developing one in the twenty first century and one that will present a number of challenges to research design. Social and therapeutic horticulture could be considered to be a form of personalised medicine insofar as a level of empathy is required in order to engage with the intervention and to maintain compliance (and possibly also to maximise the benefits). This does not mean that STH is necessarily an inappropriate intervention for those who dislike gardening or being outdoors (or who express such views initially), but including such individuals in a study will not necessarily make it a fairer test of effectiveness.

Additionally, not all of the projects that have been identified as suitable for inclusion in a trial of STH are able to take part in the randomisation process since they cannot offer control places (as outlined in Section 8) because of their particular operating procedures. Therefore, in order to conduct a viable study within a reasonable time span (which is often a requirement for funding bodies) a non-randomised arm can be included in the study and the protocol set out in Section 10 proposes such a design.

Summary and Recommendations

- Garden projects receive referrals from multiple sources and many also accept self referrals. The involvement of a large proportion of these agencies and organisations in the process of randomisation would be necessary for a successful RCT. This is unlikely to be achieved.
- Randomisation could be carried out at the point of entry into a garden project into a waiting list control and an intervention group.
- A trial of STH can include a non-randomised part of the protocol (see Section 10).

10 Conclusions and Recommendations

A protocol for a trial of social and therapeutic horticulture

In this report we have explored whether it is possible to fulfil the conditions necessary for a randomised controlled trial of social and therapeutic horticulture (STH). We have shown that there is a perceived need for such a rigorous trial of STH among practitioners, researchers and health professionals and that most of the conditions for an RCT could be fulfilled. There is also no reason why a trial of STH could be rejected as 'inappropriate' from a theoretical or scientific point of view. Although some of the issues still remain problematic, for example, that of randomisation, a protocol can be devised which represents the best possible 'fair test' of STH in the current circumstances in the UK. This is not an RCT, but just as in the evolution of the RCT in medicine, each successfully completed trial of STH represents a step towards the creation of a complete methodology. It is important to remember that the RCT method in medicine evolved over time and that many of the early trials would be subject to much criticism were exactly the same procedures employed today. However, in their time they represented major advances in both methodology and in the understanding of therapeutics. The fact that RCT methodology *exists* does not mean that it can simply be transferred *in one step* to be used to assess STH and other complex interventions. This is certainly the experience with other approaches, such as the use of therapeutic communities for people with personality disorders and mental ill health. There is still discussion and debate regarding both the value of RCTs in the field and the practicalities of conducting such trials (see Manning, 2004).

The application of rigorous scientific methodology to STH will take a number of steps and trials (and time), but this is a necessary and expected phase in the development of the evidence base underpinning STH. The effects of STH are subtle and multifactorial and vary from person to person. They are not dramatic such as those of the major drugs like penicillin or insulin, therefore, no one trial will likely to have an immediate and dramatic effect on policy or

referral. However, a successful trial will provide a direction and impetus for additional high quality research. Knowing that the evidence exists and can be investigated is an important dimension in any field of research. Just as Roger Ulrich's *View through a window* paper (Ulrich, 1984) was a catalyst to research in the field of open space and health, a successfully completed trial of STH, even if it is not an RCT, will act as a focus for further work. It will attract more research in the area, more interest from practitioners, health professionals and policy makers and ultimately more funding as the area expands (in both research and practice as in the case of open space and health).

The observations and findings contained in this report have been summarised to form the protocol shown below. The proposed study has two parts: an initial randomised controlled section of three months duration which utilises a waiting list control; this is followed by a longitudinal open section, lasting nine months, in which all participants engage in a programme of social and therapeutic horticulture.

A diagrammatic representation of the treatment and assessment protocol is provided in Figure 10.1 at the end of this section.

An Outline Protocol for a Study of Social and Therapeutic Horticulture

1. Garden projects will be included in the research if they provide a service as outlined in Section 8, i.e. a paradigm of social and therapeutic horticulture that is identifiable and distinct from hospital-based occupational therapy or which has an emphasis on production (as in the case of care farms) or which contains a dimension of formal psychotherapy (as in the case of therapeutic communities).
2. Research participants will be new starters at STH garden projects who have mental health problems and who conform to the inclusion and exclusion criteria shown in Section 7. They will have been referred by health trusts, GPs, charitable organisations (or other agencies) or have self-referred and will occupy places usually available at the project.
3. Garden staff will be responsible for explaining the study and obtaining informed consent using the standard consent forms and participant information sheets provided by the researchers.
4. Participants who agree to take part in the study will be allocated at random to one of two groups (this will be clearly explained to them), the first group will start immediately at the garden (or after any waiting period normally used at the garden); the second group will start three months later (this will be the waiting control group). Randomisation is therefore carried out after the point of entry into an STH garden project (see Section 9).
5. Clients who do not wish to enter the randomised arm of the study will be offered the opportunity to join an open longitudinal study, whereby they start immediately at the project and are assessed in the same way as the intervention group. These data will be analysed separately to the randomised part of the study (see points 13 – 15 on data analysis,

below).

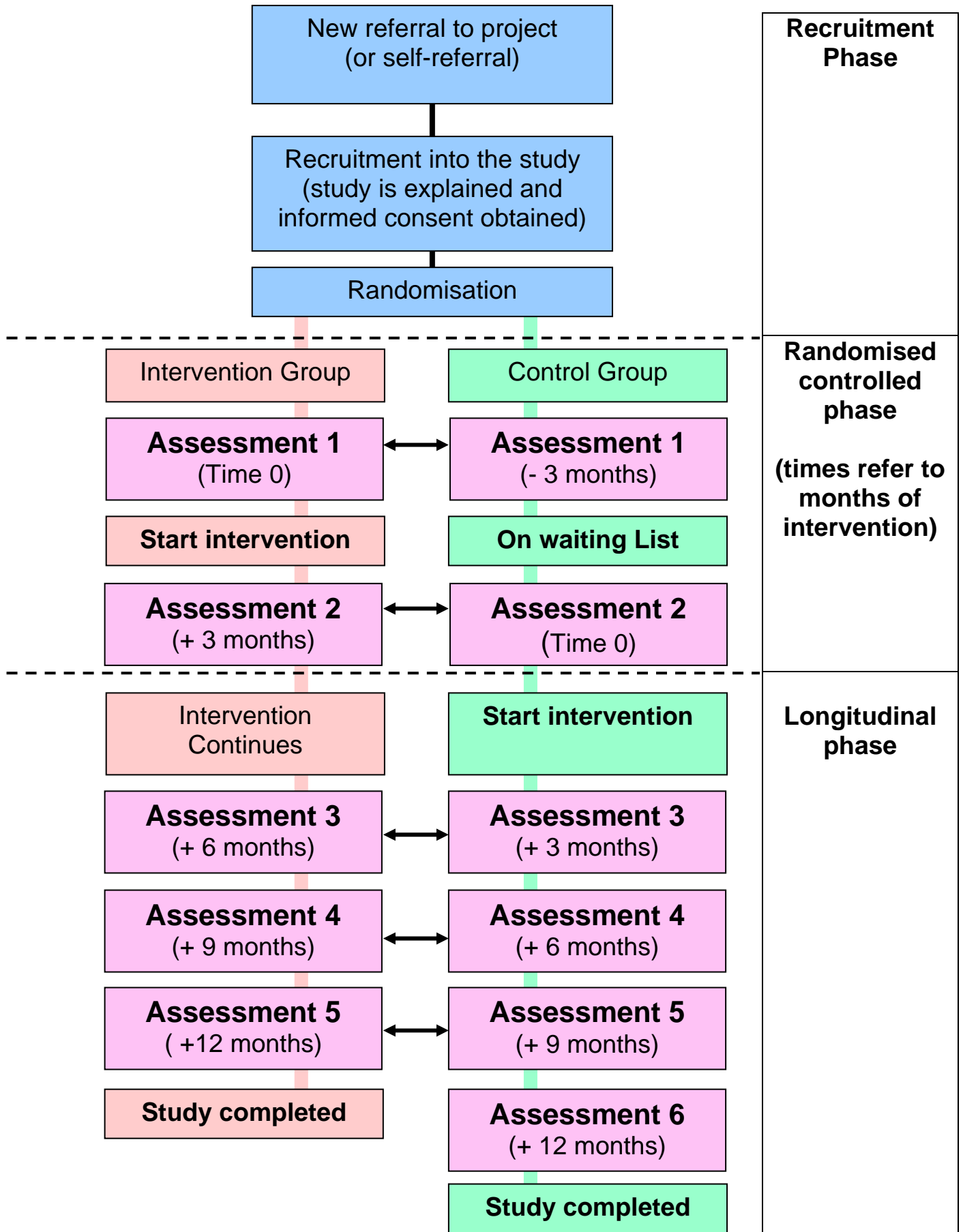
6. Garden projects that are unable to enter clients into the randomised arm of the study will enter clients into the open longitudinal study (as in 5 above).
7. Study participants will attend a project for at least two days per week and a record of attendance will be kept by project staff.
8. Study participants will take part in the standard range of activities of the project (see Section 8).
9. Assessments will be carried out using the battery of outcome measures described in Section 5. These will be conducted as follows:
 - For the intervention group**
 - at baseline (on starting the intervention); then at 3, 6, 9 and 12 months
 - For the control group**
 - at entry onto waiting list (i.e. 3 months prior to starting the intervention); at baseline (on starting the intervention); then at 3, 6, 9 and 12 months
10. Participants in the intervention group will be followed for 12 months and those in the control group for 15 months (12 month intervention period + 3 months waiting control).
11. Participants in the study will be free to remain at the garden project after the 12 month intervention period. This is in accordance with the usual operating principles of those projects i.e. study participants will be clients of garden projects who will have been offered places at the projects that are not subject to any time constraints or limits.
12. Assessments will be carried out by garden project staff and will be anonymised at the point of data collection and then posted to the

researchers. All identification of questionnaires will be by reference number alone. Research staff will be responsible for informing garden staff of impending assessments.

13. Training in the administration of the outcome measures will be provided for the garden project staff. The researchers will visit projects regularly (and as requested by staff) to offer training, advice and information to staff and participants.
14. Data analysis – randomised arm: data from participants entered into the randomised arm will be analysed to see whether there are differences in the outcome measure scores between the control and intervention group at baseline and at the assessment points during the intervention. The data will also be tested for changes in each group from baseline to intervention assessments.
15. Data analysis – open study: data from participants entered into the open study will be analysed to test for changes in outcome measure scores compared with baseline at the various assessment points; and also at follow up (six months after completion of the intervention).
16. Data analysis – pooled longitudinal data: data for all of the participants will be pooled and analysed for changes in scores compared to baseline. Data from the control group will be included in the analysis from that group's second assessment i.e. at the point of commencement of intervention (Time '0').
17. A sample of 30 participants (10 from each of the control and intervention groups of the randomised arm of the study and 10 from the longitudinal study) will be interviewed at the time of each assessment in order to explore their perceptions of their health and well-being and of the effects of attending the project or of being on the waiting list. These interviews will be recorded and subsequently analysed thematically.

18. Each garden project will recruit between 5 – 10 participants over a period of two years.
19. One hundred participants (50 per group) will be recruited to the randomisation arm of the study as recommended in Section 5. An additional number of participants will enter the open part of the study although the likely number of such participants is unknown. Those entering the open part of the study will not compromise the number of places available for the randomisation arm – these will be project clients who would have started at the project in any case.
20. Participants who drop out of the study will be followed up and interviewed to ascertain their reason for leaving. Any data collected from these participants will be included in the analysis according to the principle of ‘intention to treat’.
21. Reporting of findings, data, methodology (including randomisation procedure) and difficulties encountered will be as full as possible to ensure maximum transparency to the study, although not in a manner to compromise confidentiality of any participants.
22. Dissemination of the findings will be through a final report, a study website, journal articles and articles for professional magazines such as *Growth Point* (Thrive),

Figure 10.1: Timeline for a trial of social and therapeutic horticulture



11 References

Amberson, J.B., McMahon, B.T. and Pinner, M. (1931) 'A clinical trial of sanocrysin in pulmonary tuberculosis', *American Review of Tuberculosis*, vol. 24, pp. 401-435.

Bandura, A. (1994) 'Self-Efficacy', in: V.S. Ramachaudran (ed), *Encyclopedia of Human Behaviour*, vol. 4, pp. 71-81.

Banting, F.G., Best, C.H., Collip, J.B., Campbell, W.R. and Fletcher, A.A. (1922) 'Pancreatic extracts in the treatment of diabetes mellitus', *Canadian Medical Association Journal*, vol. 12, pp. 141-146.

Becker, S., Bryman, A. and Sempik, J. (2006) *Defining 'Quality' in Social Policy Research: Views, Perceptions and a Framework for Discussion*, Lavenham: Social Policy Association.

Berget, B. (2006) *Animal-Assisted therapy: Effects on Persons with Psychiatric Disorders Working with Farm Animals*, PhD Thesis, Aas: Norwegian University of Life Sciences and Oslo: University of Oslo.

Bradley, K.A., McDonell, M.B., Bush, K., Kivlahan, K.R., Diehr, P. and Fihn, S.D. (1998) 'The AUDIT Alcohol Consumption Questions: Reliability, Validity, and Responsiveness to Change in Older Male Primary Care Patients', *Alcoholism: Clinical and Experimental Research*, vol.22, 8 pp 1842 - 1849. [http://whqlibdoc.who.int/hq/2001/WHO_MSD_MSB_01.6a.pdf]

Brown, C.A. and Lilford, R.J. (2006) The stepped wedge design: a systematic review, *BMC Medical Research methodology*, vol.6, 54. [<http://www.biomedcentral.com/content/pdf/1471-2288-6-54.pdf>]

Burchardt, T., Le Grand, J. and Piachaud, D. (2002) 'Degrees of exclusion: Developing a dynamic, multidimensional measure', in: J. Hills, J. Le Grand,

and D. Piachaud (eds), *Understanding Social Exclusion*, New York: Oxford University Press, pp. 30-43.

Chalmers, I. (2003) 'Trying to do more good than harm in policy and practice: The role of rigorous, transparent, up-to-date evaluations', *Annals of the American Academy, AAPSS*, vol. 589, pp. 22-40.

Chalmers, T.C., Celano, P., Sacks, H.S., and Smith, H. (1983) 'Bias in treatment assignment in controlled clinical trials', *New England Journal of Medicine*, vol. 309, pp. 1358-1361.

Connell, J., Barkham, M., Stiles, W.B., Twigg, E., Singleton, N., Evans, O. and Miles, J.N.V. (2007) 'Distribution of CORE-OM scores in a general population, clinical cut-off points and comparison with the CIS-R', *British Journal of Psychiatry*, vol. 190, pp. 69-74.

Crawford, J.R., Henry, J.D. Crombie, C. and Taylor, E.P. (2001) 'Normative data for the HADS from a large non-clinical sample', *British Journal of Clinical Psychology*, vol. 40, pp. 429-434.

CSG (Core System Group) (1998) *CORE System (Information Management) Handbook*, Leeds: Core System Group.

Dehue, T. (2001) 'Establishing the experimental society: The historical origin of social experimentation according to the randomized controlled design', *American Journal of Psychology*, vol. 114, 2, pp. 283-302.

DiClemente, C.C., Prochaska, J.O. and Gibertini, M. (1985) 'Self-efficacy and the stages of self-change of smoking', *Cognitive Therapy and Research*, vol. 9, 2, pp. 181-200.

Diehr, P.D., Derleth, A.M., McKenna, S.P., Martin, M.L., Bushnell, D.M., Simon, G. and Patrick, D.L. (2006) 'Synchrony of change in depressive

symptoms, health status, and quality of life in persons with clinical depression', *Health and Quality of Life Outcomes*, vol. 4, 1, 27.

[<http://www.hqlo.com/content/pdf/1477-7525-4-27.pdf>]

Doll, R. (1998) 'Controlled trials: the 1948 watershed', *British Medical Journal*, vol. 317, pp. 1217-1220.

Edial commentary (2007) 'Fair tests of treatments in health care', The James Lind Library (www.jameslindlibrary.org), (accessed 5 June 2007).

Ernst, E. (2001) 'Commentary: Science friction - complementary/alternative medicine on the stony road from opinion to evidence', *International Journal of Epidemiology*, vol. 30, pp. 531-532.

Evans, C., Margison, F. and Barkham, M. (1998) 'The contribution of reliable and clinically significant change methods to evidence-based mental health', *Evidence Based Mental Health* (<http://ebmh.bmj.com/cgi/reprint/1/3/70>), vol. 1, pp. 70-72.

Frumkin, H. (2004) 'White coats, green plants: Clinical epidemiology meets horticulture', in: D. Relf (ed), *Expanding Roles for Horticulture in Improving Human Well-Being and Life Quality: A Proceedings of the XXVI International Horticultural Congress, Acta Horticulturae*, vol. 639.

Frumkin, H. (2005) 'Trees and human health: building the evidence base', in: Christos Th. Gallis (ed), *1st European COST E39 Conference, Forests, Trees, and Human Well-Being*, Thessalonika: Medical and Scientific Publishers.

Herrmann, C. (1997) 'International experiences with the Hospital Anxiety and Depression Scale - a review of validation data and clinical results', *Journal of Psychosomatic Research*, vol. 42, pp. 17-41.

Higgins, J.P.T, and Green, S., eds (2006) *Cochrane Handbook for Systematic Reviews of Interventions*, 4.2.6, in: The Cochrane Library, Issue 4, 2006. Chichester, UK: John Wiley & Sons, Ltd.

Hodgson, R., Alwyn, T., John, B., Thom, B. and Smith A. (2002) 'The FAST alcohol screening test', *Alcohol and Alcoholism*, vol. 37, pp. 61-66.

Hollis, S. and Campbell, F. (1999) 'What is meant by intention to treat analysis? Survey of published randomised controlled trials', *British Medical Journal*, vol. 319, pp. 670-674.

[<http://www.bmj.com/cgi/content/full/319/7211/670>]

Hróbjartsson, A., Gøtzsche, P.C. and Gluud, C. (1998) 'The controlled clinical trial turns 100 years: Fibiger's trial of serum treatment of diphtheria', *British Medical Journal*, vol. 317, pp. 1243-1245.

Jacobson, N.S. and Truax, P. (1991) 'Clinical significance: a statistical approach to defining meaningful changes in psychotherapy research', *Journal of Consulting and Clinical Psychology*, vol. 59, pp. 12-19.

Jadad, A.R. (1998) *Randomised Controlled Trials*, London: BMJ Books.

Jadad, A.R., Moore, R.A., Carroll, D., Jenkinson, C., Reynolds, D.J.M., Gavaghan, D.J. and McQuay, H.J. (1996) 'Assessing the quality of reports of randomized clinical trials: Is blinding necessary?', *Controlled Clinical Trials*, vol. 17, pp. 1-12.

Jenkinson, C., Stewart-Brown, S., Petersen, S. and Paice, C. (1999) 'Assessment of the SF-36 version 2 in the United Kingdom', *Journal of Epidemiology and Community Health*, vol. 53, pp. 46-50.

Kunz, R. and Oxman, A.D. (1998) 'The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials', *British Medical Journal*, vol. 317, pp. 1185-1190.

Lenth, R. V. (2006) *Java Applets for Power and Sample Size* [Computer software], retrieved July 2006, from <http://www.stat.uiowa.edu/~rlenth/Power>.

Lind, J. (1753) *A Treatise of the Scurvy. In three parts. Containing an inquiry into the nature, causes and cure, of that disease. Together with a critical and chronological view of what has been published on the subject*, Edinburgh: Printed by Sands, Murray and Cochran for A Kincaid and A Donaldson.
[Facsimile available at:
http://www.jameslindlibrary.org/trial_records/17th_18th_Century/lind/lind_kp.html] (accessed 5 June 2007).

Linde, K., Jonas, W.B., Melchart, D. and Willich, S. (2001) 'The methodological quality of randomized controlled trials of homeopathy, herbal medicines and acupuncture', *International Journal of Epidemiology*, vol. 30, pp. 526-531.

Lister, J. (1870) 'Effects of the antiseptic system of treatment upon the salubrity of a surgical hospital', *Lancet*, vol. 1, pp. 40-42.

Manning, N. (2004) 'The Gold Standard, What are RCTs and Where Did They Come From?', in: Jan Lees, Nick Manning, Diana Menzies and Nicola Morant (eds), *A Culture of Enquiry: Research Evidence and the Therapeutic Community*, London: Jessica Kingsley Publishers.

Medical Research Council (1948) 'Streptomycin treatment of pulmonary tuberculosis', *British Medical Journal*, vol. 2, pp. 769-782.

Milgrom, L.R. (2005) 'Are randomized controlled trials (RCTs) redundant for testing the efficacy of homeopathy? A critique of RCT methodology based on entanglement theory', *The Journal of Alternative and Complementary Medicine*, vol. 11, 5, pp. 831-838.

Milligan, C., Bingley, A. and Gatrell, A. (2003) *Cultivating Health: A Study of Health and Mental Well-being Amongst Older People in Northern England*, Lancaster: Institute for Health Research, Lancaster University.

Oakley, A. (1998) 'Experimentation and social interventions: a forgotten but important history', *British Medical Journal*, vol. 317, pp. 1239-1242.

Schulz, K. F. (2000) 'Assessing allocation concealment and blinding in randomised controlled trials: why bother?', *Evidence-Based Medicine*, vol. 5. (March/April), 38.

[<http://ebm.bmj.com/cgi/content/full/5/2/38>]

Schulz, K.F., Chalmers, I., Hayes, R.J. and Altman, D. (1995) 'Empirical evidence of bias', *Journal of the American Medical Association*, vol. 273, 5, pp. 408-412.

Schwarzer, R. (2007) *General Perceived Self-Efficacy Scale*, [Data file] retrieved from, <http://userpage.fu-berlin.de/~health/selfscal.htm>.

Schwarzer, R., and Jerusalem, M. (1995) 'Generalized Self-Efficacy scale', in: J. Weinman, S. Wright, & M. Johnston (eds), *Measures in Health Psychology: A User's Portfolio. Causal and Control Beliefs*, pp. 35-37, Windsor: NFER-NELSON.

Sempik, J., Aldridge, J. and Becker, S. (2003) *Social and Therapeutic Horticulture: Evidence and Messages from Research*, Reading: Thrive and Loughborough: CCFR.

Sempik, J., Aldridge, J. and Becker, S. (2005) *Health, Well-being and Social Inclusion, Therapeutic Horticulture in the UK*, Bristol: The Policy Press.

Sempik, J. and Spurgeon, T. (2006) 'Lessons learnt – Evidence from practice: The use of plants and horticulture in promoting health and well-being' in: *Proceedings of the 6th International Congress on Education in Botanic*

Gardens, Richmond: Botanic Gardens Conservation International in association with Oxford: University of Oxford Botanic Garden (ISBN 1-905164-13-0).

Skevington, S.M., Lotfy, M. and O'Connell, K.A. (2004) 'The World Health Organization's WHOQOL-BREF quality of life assessment: Psychometric properties and results of the international field trial A Report from the WHOQOL Group', *Quality of Life Research*, vol. 13, pp. 299-310.

Smith, G.C.S. and Pell, J.P. (2003) 'Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials', *BMJ*, vol. 327, pp. 1459-1461.

Snaith, R.P. and Zigmond, A.S. (1994) *HADS: Hospital Anxiety and Depression Scale*, Windsor: NFER Nelson.

Tröhler, U . (2003) *James Lind and scurvy, 1747 to 1795*, The James Lind Library, (www.jameslindlibrary.org), (accessed Tuesday 3 October 2006).

Ulrich, R.S (1984) 'View through a window may influence recovery from surgery', *Science*, vol. 224, pp. 420-421.

Zigmond, A.S. and Snaith, R.P. (1983) 'The Hospital Anxiety and Depression Scale', *Acta Psychiatr Scand*, vol. 67, pp. 361-370.

Appendix

Outcome measures used in this study

1. Clinical Outcomes in Routine Evaluation (CORE) Outcome Measure
2. Hospital Anxiety and Depression Scale (HADS)
3. General Self Efficacy Scale
4. Use of Services and Substances
5. WHOQOL-BREF – for copyright reasons it is not possible to reproduce this form here. It may be downloaded from the WHO Field Centre for the Study of Quality of Life, University of Bath after online registration with the centre, website: [<http://www.bath.ac.uk/whoqol/questionnaires/info.cfm>]

**1. Clinical Outcomes in Routine Evaluation (CORE)
Outcome Measure**

CLINICAL
OUTCOMES in
ROUTINE
EVALUATION

**OUTCOME
MEASURE**

Site ID	<input type="text"/>	<input type="text"/>	Male	<input type="checkbox"/>
letters only	<input type="text"/>	numbers only	Age	Female
Client ID	<input type="text"/>	<input type="text"/>	Stage Completed	Stage
Therapist ID	<input type="text"/>	numbers only (1)	S Screening	<input type="checkbox"/>
Sub codes	<input type="text"/>	numbers only (2)	R Referral	
Date form given	<input type="text"/>	<input type="text"/>	A Assessment	
	<input type="text"/>	<input type="text"/>	F First Therapy Session	
	<input type="text"/>	<input type="text"/>	P Pre-therapy (unspecified)	
	<input type="text"/>	<input type="text"/>	D During Therapy	
	<input type="text"/>	<input type="text"/>	L Last therapy session	Episode
	<input type="text"/>	<input type="text"/>	X Follow up 1	<input type="checkbox"/>
	<input type="text"/>	<input type="text"/>	Y Follow up 2	

IMPORTANT - PLEASE READ THIS FIRST

This form has 34 statements about how you have been **OVER THE LAST WEEK**.
Please read each statement and think how often you felt that way last week.
Then tick the box which is closest to this.
Please use a dark pen (not pencil) and tick clearly within the boxes.

Over the last week	Not at all	Only Occasionally	Sometimes	Often	Most or all the time	OFFICE USE ONLY
1 I have felt terribly alone and isolated	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> F
2 I have felt tense, anxious or nervous	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> P
3 I have felt I have someone to turn to for support when needed	<input type="checkbox"/> 4	<input type="checkbox"/> 3	<input type="checkbox"/> 2	<input type="checkbox"/> 1	<input type="checkbox"/> 0	<input type="checkbox"/> F
4 I have felt O.K. about myself	<input type="checkbox"/> 4	<input type="checkbox"/> 3	<input type="checkbox"/> 2	<input type="checkbox"/> 1	<input type="checkbox"/> 0	<input type="checkbox"/> W
5 I have felt totally lacking in energy and enthusiasm	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> P
6 I have been physically violent to others	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> R
7 I have felt able to cope when things go wrong	<input type="checkbox"/> 4	<input type="checkbox"/> 3	<input type="checkbox"/> 2	<input type="checkbox"/> 1	<input type="checkbox"/> 0	<input type="checkbox"/> F
8 I have been troubled by aches, pains or other physical problems	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> P
9 I have thought of hurting myself	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> R
10 Talking to people has felt too much for me	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> F
11 Tension and anxiety have prevented me doing important things	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> P
12 I have been happy with the things I have done.	<input type="checkbox"/> 4	<input type="checkbox"/> 3	<input type="checkbox"/> 2	<input type="checkbox"/> 1	<input type="checkbox"/> 0	<input type="checkbox"/> F
13 I have been disturbed by unwanted thoughts and feelings	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> P
14 I have felt like crying	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> W

Please turn over

Over the last week

Not at all Only Occasionally Sometimes Often Most or all the time OFFICE USE ONLY

15	I have felt panic or terror	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> P
16	I made plans to end my life	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> R
17	I have felt overwhelmed by my problems	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> W
18	I have had difficulty getting to sleep or staying asleep	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> P
19	I have felt warmth or affection for someone	<input type="checkbox"/> 4	<input type="checkbox"/> 3	<input type="checkbox"/> 2	<input type="checkbox"/> 1	<input type="checkbox"/> 0	<input type="checkbox"/> F
20	My problems have been impossible to put to one side	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> P
21	I have been able to do most things I needed to	<input type="checkbox"/> 4	<input type="checkbox"/> 3	<input type="checkbox"/> 2	<input type="checkbox"/> 1	<input type="checkbox"/> 0	<input type="checkbox"/> F
22	I have threatened or intimidated another person	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> R
23	I have felt despairing or hopeless	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> P
24	I have thought it would be better if I were dead	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> R
25	I have felt criticised by other people	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> F
26	I have thought I have no friends	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> F
27	I have felt unhappy	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> P
28	Unwanted images or memories have been distressing me	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> P
29	I have been irritable when with other people	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> F
30	I have thought I am to blame for my problems and difficulties	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> P
31	I have felt optimistic about my future	<input type="checkbox"/> 4	<input type="checkbox"/> 3	<input type="checkbox"/> 2	<input type="checkbox"/> 1	<input type="checkbox"/> 0	<input type="checkbox"/> W
32	I have achieved the things I wanted to	<input type="checkbox"/> 4	<input type="checkbox"/> 3	<input type="checkbox"/> 2	<input type="checkbox"/> 1	<input type="checkbox"/> 0	<input type="checkbox"/> F
33	I have felt humiliated or shamed by other people	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> F
34	I have hurt myself physically or taken dangerous risks with my health	<input type="checkbox"/> 0	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> R

THANK YOU FOR YOUR TIME IN COMPLETING THIS QUESTIONNAIRE

Total Scores

 → →

Mean Scores

(Total score for each dimension divided by number of items completed in that dimension)

(W) (P) (F) (R) All items All minus R

2. Hospital Anxiety and Depression Scale (HADS)

Instructions: Read each item and place a firm tick in the box opposite the reply which comes closest to how **you have been feeling in the past week**. Don't take too long over your replies. Your immediate reaction to each item will probably be more accurate than a long thought out response.

I feel tense or 'wound up':	<input type="checkbox"/>
Most of the time	<input type="checkbox"/>
A lot of the time	<input type="checkbox"/>
Time to time, occasionally	<input type="checkbox"/>
Not at all	<input type="checkbox"/>

I still enjoy the things I used to enjoy:	<input type="checkbox"/>
Definitely as much	<input type="checkbox"/>
Not quite so much	<input type="checkbox"/>
Only a little	<input type="checkbox"/>
Not at all	<input type="checkbox"/>

I get a sort of frightened feeling like something awful is about to happen:	<input type="checkbox"/>
Very definitely and quite badly	<input type="checkbox"/>
Yes, but not too badly	<input type="checkbox"/>
A little, but it doesn't worry me	<input type="checkbox"/>
Not at all	<input type="checkbox"/>

I can laugh and see the funny side of things:	<input type="checkbox"/>
As much as I always could	<input type="checkbox"/>
Not quite so much now	<input type="checkbox"/>
Definitely not so much now	<input type="checkbox"/>
Not at all	<input type="checkbox"/>

Worrying thoughts go through my mind:	<input type="checkbox"/>
A great deal of the time	<input type="checkbox"/>
A lot of the time	<input type="checkbox"/>
From time to time but not too often	<input type="checkbox"/>
Only occasionally	<input type="checkbox"/>

I feel cheerful:

	6
Not at all	<input type="checkbox"/>
Not often	<input type="checkbox"/>
Sometimes	<input type="checkbox"/>
Most of the time	<input type="checkbox"/>

I can sit at ease and feel relaxed:

	7
Definitely	<input type="checkbox"/>
Usually	<input type="checkbox"/>
Not often	<input type="checkbox"/>
Not at all	<input type="checkbox"/>

I feel as if I am slowed down:

	8
Nearly all of the time	<input type="checkbox"/>
Very often	<input type="checkbox"/>
Sometimes	<input type="checkbox"/>
Not at all	<input type="checkbox"/>

I get a sort of frightened feeling like 'butterflies in the stomach':

	9
Not at all	<input type="checkbox"/>
Occasionally	<input type="checkbox"/>
Quite often	<input type="checkbox"/>
Very often	<input type="checkbox"/>

I have lost interest in my appearance:

	10
Definitely	<input type="checkbox"/>
I don't take as much care as I should	<input type="checkbox"/>
I may not take quite as much care	<input type="checkbox"/>
I take just as much care as ever	<input type="checkbox"/>

I feel restless as if I have to be on the move: **11**

Very much indeed

Quite a lot

Not very much

Not at all

I look forward with enjoyment to things: **12**

A much as I ever did

Rather less than I used to

Definitely less than I used to

Hardly at all

I get sudden feelings of panic: **13**

Very often indeed

Quite often

Not very often

Not at all

I can enjoy a good book or radio or TV programme: **14**

Often

Sometimes

Not often

Very seldom

3. General Self Efficacy Scale

Instructions: Read each item and place a firm tick in the box opposite which comes closest to how you feel about the statement.

1 I can always manage to solve difficult problems if I try hard enough.	Not at all true	<input type="checkbox"/>
	Hardly true	<input type="checkbox"/>
	Moderately true	<input type="checkbox"/>
	Exactly true	<input type="checkbox"/>

2 If someone opposes me, I can find the means and ways to get what I want.	Not at all true	<input type="checkbox"/>
	Hardly true	<input type="checkbox"/>
	Moderately true	<input type="checkbox"/>
	Exactly true	<input type="checkbox"/>

3 It is easy for me to stick to my aims and accomplish my goals.	Not at all true	<input type="checkbox"/>
	Hardly true	<input type="checkbox"/>
	Moderately true	<input type="checkbox"/>
	Exactly true	<input type="checkbox"/>

4 I am confident that I could deal efficiently with unexpected events.	Not at all true	<input type="checkbox"/>
	Hardly true	<input type="checkbox"/>
	Moderately true	<input type="checkbox"/>
	Exactly true	<input type="checkbox"/>

5 Thanks to my resourcefulness, I know how to handle unforeseen situations.	Not at all true	<input type="checkbox"/>
	Hardly true	<input type="checkbox"/>
	Moderately true	<input type="checkbox"/>
	Exactly true	<input type="checkbox"/>

6 I can solve most problems if I invest the necessary effort.

Not at all true	<input type="checkbox"/>
Hardly true	<input type="checkbox"/>
Moderately true	<input type="checkbox"/>
Exactly true	<input type="checkbox"/>

7 I can remain calm when facing difficulties because I can rely on my coping abilities.

Not at all true	<input type="checkbox"/>
Hardly true	<input type="checkbox"/>
Moderately true	<input type="checkbox"/>
Exactly true	<input type="checkbox"/>

8 When I am confronted with a problem, I can usually find several solutions.

Not at all true	<input type="checkbox"/>
Hardly true	<input type="checkbox"/>
Moderately true	<input type="checkbox"/>
Exactly true	<input type="checkbox"/>

9 If I am in trouble, I can usually think of a solution.

Not at all true	<input type="checkbox"/>
Hardly true	<input type="checkbox"/>
Moderately true	<input type="checkbox"/>
Exactly true	<input type="checkbox"/>

10 I can usually handle whatever comes my way.

Not at all true	<input type="checkbox"/>
Hardly true	<input type="checkbox"/>
Moderately true	<input type="checkbox"/>
Exactly true	<input type="checkbox"/>

4. Use of Services and Substances

Use of Services

Please answer each question by placing a firm tick in the spaces below.

1. We would like to ask you how many of the services listed below you have used in the last **SIX MONTHS**. In the last six months, how often:

	No	Once	2 or 3 times	More than 3 times	How many times days
Have you seen your GP?					
Have you had to make an emergency appointment to see your GP?					
Have you attended an Accident and Emergency department?					
Had an admission to hospital? (If yes, please note the number of admissions <i>and</i> total number of days)					
Have you seen a social worker, benefits or housing worker?					
Have you seen a psychiatrist, community psychiatric nurse or psychologist? (including unplanned visits)					
Have you lost time from work due to ill health?					
Have you had contact with the police?					
Have you been arrested?					
Have you been charged with an offence?					

2. We would like to ask you a few questions about your use of alcohol and other drugs in the **LAST THREE MONTHS**.

All the answers that you give will be held in strict confidence. You will not be identified and the information will not be passed on to anyone else. Please answer each question by placing a firm tick in the spaces provided.

	Never	Less than monthly	Monthly	Weekly	Daily or almost daily
<p>for MEN: How often do you have EIGHT or more drinks on one occasion?</p> <p>for WOMEN: How often do you have SIX or more drinks on one occasion?</p> <p><i>(One drink is ½ pint of beer or 1 glass of wine or 1 single spirit measure)</i></p>					
How often during the <u>last year</u> have you been unable to remember what happened the night before because you had been drinking?					
How often during the <u>last year</u> have you failed to do what is normally expected from you because of drinking? <i>(For example, missing an appointment or not going to work)</i>					
In the <u>last year</u> has a relative or friend, or a doctor or other health worker been concerned about your drinking or suggested you cut down?	Please answer either:				
	Yes (on more than one occasion)				
	Yes (on one occasion)				
	No				
How often, if ever, have you used cannabis in the <u>last three months</u> ?					
Please write the name of any other drugs you have used in the <u>last three months</u> ?					
How often, have you used these drugs in the <u>last three months</u> ?					



The Path of Life garden was conceived and constructed by a group of people recovering from mental health problems. The concept is that of a path which is reflective of peoples' journey through their illness. The start of the path is bleak and barren, gradually softening and eventually developing into a serene and inspirational setting.



Mental Health Foundation