Summary

The study "Options for treatment outcome research in the 'terbeschikkingstelling' (tbs)", was aimed at: (1) listing research methods and designs for efficacy and effectiveness research and (2) describing whether these methods are usable for investigating mechanisms of change and responsivity research. The tbs order is a penal measure. It can be imposed by the court upon mentally disordered offenders who are considered not to be responsible or to have diminished responsibility for their offence. Effectiveness and responsivity research will yield the evidence base for tbs treatment guidelines, composed of empirically supported therapies. Guidelines include descriptions of what works best for whom under which circumstances. The present study yields a realistic estimate of possibilities for and limitations to treatment outcome research in the tbs.

This report comprises three sections. The first section specifies the research question and our approach to answering it (Chapter 1). In addition, the tbs setting in which the treatment outcome research will be conducted is described in words and numbers (Chapter 2). Tbs patients are a quite heterogeneous population regarding their criminal history as well as the psychopathology they suffer from. The psychopathology is often severe and complex, as evidenced by high comorbidity rates in most tbs patients.

The second section (Chapter 3 and 4) considers the background of evidence-based treatment. The evidence-based treatment concept stems from the medical field, first moved to the mental health profession, and more recently to the tbs setting. Three steps need to be taken before the tbs setting can claim to work according to evidence-based principles: (1) scientific research to obtain the empirical support for treatments, (2) formulation of treatment guidelines based on empirical data, and (3) implementation and dissemination of these guidelines in the field. Thus, for the tbs setting, a first requirement is to build an evidence base for specific interventions and treatment programs through scientific research. Which of these are efficacious and effective? The research methods described in this report are primarily aimed at evaluation of these interventions and treatments. The suitability of treatment outcome research to address questions regarding the effectiveness of the tbs method as a whole, including non-therapeutic measures such as appropriate selection for treatment termination and posttreatment services that are offered, is limited.

An important issue in treatment outcome research is the choice of a proper outcome measure. Regarding outcome measures for effectiveness research in the tbs, the following is concluded: although reduction of actual recidivism may be the primary goal of tbs treatment, this variable is ill-suited and of limited value as an outcome measure of controlled effectiveness research, as it is in time far removed from the evaluated intervention and may be determined by many other factors than the treatment alone. In addition, recidivism is a rather insensitive outcome measure, necessitating a high number of patients participating in a trial to ensure sufficient statistical power. Better outcome variables are the extent to which specific treatments goals have been attained or the reduction of *risk* for recidivism.

The third section of this report describes methods of research to evaluate treatment effect. Chapter 5 describes methodological characteristics of several study designs for treatment outcome research. Issues related to internal and external validity are discussed and we address the difference between demonstrating efficacy (Is treatment X superior to control treatment Y?) under highly controlled circumstances and effectiveness of treatments (Is treatment efficacious under less controlled circumstances and is the treatment sufficient?) in clinical practice. An important feature of a study design is whether it can be informative regarding the influence of responsivity factors. Besides these defining characteristics of study designs, each type of study comes with specific requirements for its implementation in clinical practice. Several of these requirements are described, such as randomisation, placebo treatment, and statistical power. These requirements have to be taken into account when we decide on the feasibility of study designs for the tbs.

Chapter 6 presents three types of outcome research: *controlled research* with randomised controlled trials, *outcome monitoring* and *single case designs*. Randomised controlled trials are the gold standard for controlled treatment outcome research, and the method of preference when it comes to demonstrating the efficacy of interventions (i.e., the outcome for experimentally-treated patients as compared to no treatment controls). In its simplest form, an RCT implies that patients are randomly assigned to either an experimental treatment condition (comprising the active treatment ingredients) or a control treatment without active ingredients (a wait list or placebo condition). Comparison of these groups indicates whether the experimental treatment is superior to the control treatment. The RCT design is a truly experimental design and offers ample possibilities to control for treats to the internal validity of the study. However, at the same time this optimized control threatens the external validity of the study: Prototypical patients and intensely supervised treatments are hard to find in everyday clinical practice. Single case designs are truly experimental studies that investigate changes within a single person. These designs are especially suited for a quick first impression on the efficacy of a new treatment. The generalizability of the findings is, however, limited until findings are replicated in a sizeable group of patients. Furthermore, single case designs are only feasible with specific interventions, namely interventions that can be "turned on and off at will". RCT's and single case designs are both less suited to investigate responsivity factors, factors that potentially influence treatment outcome. A key ingredient of the methodology of experimental designs is to try to minimize variance in responsivity factors, which renders them less appropriated for the investigation of these very factors.

Outcome monitoring aims at demonstrating the effectiveness of treatments in the clinical reality of everyday practice. Outcome monitoring implies routine data collection regarding the type and quality of treatments delivered, and simultaneous monitoring of the functioning and psychological wellbeing of the patient. At set time points, standardised outcome measures are administered. With outcome monitoring, current practice and new experimental treatments can be evaluated under controlled or uncontrolled conditions. Outcome monitoring is very well-suited for an investigation into factors related to patients' responsivity to treatment (which personality or crime characteristics predict good outcome) and which treatment conditions do so. Thus, to address the question who benefits most from therapeutic interventions in the tbs and which treatment conditions are optimal for a positive outcome this design is suitable.

The three research designs are compared regarding their methodological strengths and weaknesses and their feasibility in the tbs setting. We conclude that the implementation of an RCT, albeit the gold standard for demonstrating efficacy of an intervention, is complex because of its required randomisation and incorporation of a control group in the design. Because of power considerations, substantial numbers of patients are required, limiting the number of RCT's that can be conducted in the tbs. The internal validity of outcome monitoring is inferior to RCT's but its external validity is superior. Furthermore, outcome monitoring is very feasible for the tbs setting. Single case designs and RCT's both have a high validity for demonstrating efficacy of the evaluated treatment. Single case designs are easily implemented, but their external validity is, however, limited.

Research endeavours in the tbs should aim at generating the much-needed evidence base (What works best for whom?) and, wherever feasible, constraints to empirical research should be counteracted. The last chapter addresses these constraints by describing requirements for treatment outcome research in the tbs and providing recommendations on how to build an evidence-based foundation for treatment in the tbs. In addition, outcome monitoring is recommended as a means of enabling small scale randomised controlled trials, and single case studies. A plea is made for the creation of an infrastructure for outcome monitoring, which would be the practical implementation of outcome monitoring. The infrastructure that is required for outcome monitoring encompasses monitoring of the treatments that are delivered and monitoring of outcomes. Treatment outcome research is a complicated endeavour and, apart

from single case designs, valid results are only obtained with sufficient numbers of patients. Thus, most study designs pose requirements that go beyond the means of single tbs clinics. Collaboration among clinics should be stimulated, by striving for a common set of outcome measures. Another form of collaboration is organising RCT's primarily through multi-centre trials. A multi-centred approach could be a condition for applications for financial support for outcomes research.