Aspects of using physical training on patients with substance dependence and additional mental distress

Asgeir Mamen

Dissertation for the degree philosophiae doctor (PhD)

at the University of Bergen

2011

Dissertation date: 16th of June

Scientific environment

Department of Psychiatry, Sogn og Fjordane Central Hospital, Førde
The West Norwegian Research Education Network (VNN-FU), Senter for vitskapsteori, Bergen
Sogn og Fjordane University College, Faculty of teacher education and sport, Sogndal

Acknowledgements

”Do you hold a PhD?” The question comes from Egil Martinsen, and the scene is a meeting room at the Department of Psychiatry at Førde Central Hospital, Norway. Present are Atle Skrede, Harald Munkvold, Egil Martinsen and myself. We have been discussing how to communicate the findings of the newly finished project to the outside world. A bit bewildered, I answer that I do not have a doctoral degree as such, but the equivalent. “You see,” continues Martinsen, “I think this project deserves a PhD, so if you are willing to write it, I’ll supervise you!” After some deliberation, I accepted the invitation, and the result is found in these pages.

During the process, I have met with much support and help from many people. First of all, I would like to thank my wife Tone and my colleagues at Sogn og Fjordane University College for their support and patience. I also wish to convey my sincere respect to Atle Skrede and Harald Munkvold for their professional implementation of the project. The West Norwegian Research Education Network (VNN-FU), represented by Professor Bente Wold, proved to be a very valuable asset for me during the educational part of the process. However, I am most grateful to my supervisors, Professor Ståle Pallesen, Faculty of Psychology, University of Bergen and Professor Egil W. Martinsen, Faculty of Medicine, University of Oslo. With immaculate patience and constructive comments, they have encouraged my progress at every step. Dr. Jon Ingulf Medbø of the National Institute of Occupational Health also deserves warm thanks for giving me much useful advice on writing articles.

Asgeir Mamen

Asker/Sogndal

12.11.10

Abstract

Physical fitness and mental health were assessed in 47 persons with substance dependence and additional mental distress when they started on a training project. After an average of 11 months in the project, the 36 remaining subjects (33 with sufficient data) were similarly assessed upon discharge. Direct measurements of maximal aerobic power and a lactate profile test at inclusion served as baseline and basis for the design of individual training programmes. The maximal aerobic power was higher than that previously reported for alcohol-dependent subjects, 37 ml•kg^{-1}•min^{-1} (Paper I), but still subnormal relative to the age-matched normal healthy population. The training programme included "training partners", who trained together with the participants. The training load varied from 150 to 500 h. The increase in maximal aerobic power was a modest 5%, but lactate threshold performance improved rather more, by 11% (Paper II). Nineteen persons responded to the training, improving maximal aerobic power by 13% and lactate threshold performance by 16%. The remaining 14 did not improve in physical fitness. The only significant difference between these two groups was the percentage of training performed at high intensity above threshold heart rate, 32% for responders compared to 18% for non-responders. Measurements of mental distress in terms of anxiety, depression, social phobic symptoms, general psychiatric problems, alcohol dependence and drug addiction were administered at inclusion and discharge, revealing a substantial level of mental distress at the outset. The discharge data showed reductions in all measures of mental distress, but no significant change in measures of alcohol dependence and drug addiction. The magnitude of improvement in mental distress paralleled the improvement in physical fitness (Paper III). The study shows that substance-dependent persons are capable of performing a high volume of training for an extended time. The use of training partners, a direct test of maximal aerobic power and a lactate profile test all proved feasible for this group of patients.

List of publications

I

Mamen, A. & Martinsen, E.W. (2009). The aerobic fitness of substance abusers voluntary participating in a rehabilitation project. *Journal of Sports Medicine & Physical Fitness*, *49*, 187–93.

II

Mamen, A. & Martinsen, E.W. (2010). Development of aerobic fitness of individuals with substance abuse/dependence following long-term individual physical training. *European Journal of Sport Science, 10,* 255–262. DOI: 10.1080/17461390903377126.

III

Mamen, A., Pallesen, S., & Martinsen, E.W. (2011). Changes in mental distress following individualised physical training in patients suffering from chemical dependence. *European Journal of Sport Science*, 11, 269–276. DOI: 10.1080/17461391.2010.509889.

List of tables

List of abbreviations

95%CI ..... 95 per cent confidence interval
%HRmax ..... Per cent of maximal heart rate
ADHD ..... Attention Deficit Hyperactivity Disorder
ATP ..... Adenosin-Tri-Phosphate
AUDIT ..... Alcohol Use Disorders Identification Test
BAI ..... Beck Anxiety Inventory
BDI-2/BDI ..... Beck Depression Inventory II
BLC ..... Blood Lactate Concentration
BSPS ..... Brief Social Phobia Scale
COPD ..... Chronic Obstructed Pulmonary Disease
DAST-20/DAST ..... Drug Abuse Screening Test 20
DUDIT ..... Drug Use Disorders Identification Test
ES ..... Effect size (Cohen's d)
GP ..... General Practitioner
h/w ..... hours pr. week
H+ ..... Hydrogen ions
HR ..... Heart rate
MAP ..... Maximal Aerobic Power (=$\dot{V}O_{2max}$=maximal oxygen uptake)
MET ..... Metabolic Unit (~3.5 ml•kg^{-1}•min^{-1})
N33 ..... Post-treatment number of subjects
N47 ..... Baseline number of subjects
RER ..... Respiratory Exchange atio: $\dot{V}CO_2/\dot{V}O_2$
RPE ..... Rating of Perceived Exertion
SCL-90R ..... Symptoms Checklist 90 Revised
GSI ..... Global Severity Index
SD ..... Standard deviation
SEM ..... Standard Error of Mean
SMAST ..... Short Michigan Alcoholic Screening Test
VT1, VT2 ..... Ventilatory Threshold 1 and 2
YMCA ..... Young Mens Christian Organisation
Å/R ..... Åstrand/Ryhming submaximal cycle ergometer test

Table of contents

1 Aspects of using physical training on patients with substance dependence and additional mental distress

In this thesis, some aspects of giving physical training to patients with substance dependence[1] and additional mental distress will be addressed. There is little evidence-based knowledge on how substance-dependent people react to physical training, and how fit they are, as a group. Neither do we know much of the relationship between changes in fitness and mental distress in this group.

The two main outcome variables in this thesis are **maximal aerobic power** (MAP, also known as maximal oxygen uptake or $\dot{V}O_{2max}$ in the literature) and **lactate threshold performance**, the highest physical load one can sustain without an increase in blood lactate concentration (Weltman, 1995). Direct assessment of MAP requires maximal exertion of the subject and a metabolic analyser to analyse the expired air. MAP is often used as an indicator of physical fitness and endurance capacity, so the measurements of aerobic power in this thesis can be seen as an expression of the fitness and endurance capacity of the project participants.

Lactate threshold performance is here recorded from a submaximal load, based on measurements of blood lactate concentration from capillary blood, and is likewise a good estimator of endurance capacity (Weltman, 1995). These main variables will subsequently be compared with variables such as depression, anxiety, general psychiatric problems, social phobia and substance dependence.

1.1 Substance dependence

The use of intoxicating chemicals has a long history. Egyptian hieroglyphs and ancient Chinese writings recount the use of drugs and how to prepare them. It seems that

[1] Substance dependence=dependence on alcohol and/or intoxicating substances

alcohol use goes back to Palaeolithic times (~8000 years ago) and that the use of psychoactive substances has played a role in social life, religion and culture from the earliest human societies (Westermeyer, 2005). Today, abuse of and dependence on intoxicating substances represents a substantial societal problem worldwide.

1.2 Prevalence and co-morbidity

Alcohol and drug abuse and dependence are increasing both in Norway and the rest of Europe (Anderson & Baumberg, 2006). Although the level of alcohol intake in Norway is one of the lowest in Europe, the 2008 mean value for pure alcohol consumption by Norwegians above 15 years was 6.75 L $\bullet yr^{-1}$ (Statistics Norway, 2010), an increase from previous years. Beck, Wright, Newman, and Liese (1993) noted that in the USA, 10% of the population consume half of the total consumption of alcohol, indicating a small number of very high consumers. Although mean intake is relatively low in Norway, a minority has an unhealthy intake. This represents a health risk to themselves and requires action by the public health service. In 2008, the recorded number of hospital admittances due to substance abuse was about 12 500, representing almost 575 000 patient days. Males represented about 2/3 of the persons treated (Statistics Norway, 2010).

Together with depression and anxiety, substance abuse and dependence are the most common psychiatric disorders (The Norwegian Institute of Public Health, 2010). In Sweden, Nordström and Bodlund (2008) found that every third patient seeking treatment at a regional somatic hospital also showed signs of depression, anxiety, and/or alcohol problems. Bijl, Ravelli and van Zessen (1998) found psychiatric disorders to be common in the general population in the Netherlands. Regier, Farmer, Rae, Locke, Keith, Judd, and Goodwin (1990) determined the co-morbidity of drug-

related diagnoses and mental disorders in a sample of 20 000 persons in the USA. A total of 37% of those with an alcohol disorder had co-morbid mental disorders and 53% of the persons with drug disorders also had other mental disorders. In a review, Armstrong and Costello (2002) found that 50% to 80% of all types of substance abusers have also met the criteria for at least one other psychiatric diagnosis at some time in their lives. They also reported concurrent co-morbidity of substance abuse and depression at a median of 19%. The reasons for this co-morbidity can be many; some use substances as self-medication for psychiatric illness, while for others the use of substances will lead to psychiatric disorders. A third explanation is that there are common factors which dispose people to both disorders. Anxiety and substance abuse show lower co-morbidity than depression and substance abuse (Armstrong & Costello, 2002).

1.3 Public health finances and treatment costs related to chemical abuse and dependence

Hospitalisation due to chemical abuse and dependence represents a substantial part of the costs of the health care system in Norway. Melberg (2005) discusses the problem of quantifying the costs of drug treatment. Different methods of calculation give different results. There is no gold standard for these calculations, and the results will depend on the choice of factors entered into them. Thus, all cost estimates will be subjective (Melberg, 2005). In the USA, Russo and Elixhauser (2003) found that alcohol was involved in 3% of all hospital stays.

Cognitive behavioural therapy (CBT), motivational interviewing (MI) and the 12-step model of the Alcoholics Anonymous tradition are well-documented forms of psychological treatment (Frances, Miller, & Mack, 2005). Medication plays an

important role in the management of opiate dependence, but has limited use in the treatment of alcohol-related disorders (ibid.). Therapy resources are limited, and many patients do not get access to evidence-based forms of therapy. It is therefore of interest to explore other ways of treatment, particularly those which are easily available and inexpensive. One such treatment could be physical training, which has a low cost and can be given to both outpatients and inpatients (Skrede, Munkvold, Watne, & Martinsen, 2006).

1.4 Physical training and cardio-respiratory fitness

Physical training is defined as the systematic use of physical activity to improve or maintain physical performance level/health (Skaset, 1976). Physical training encompasses endurance, sprint/speed, strength, flexibility and technical training. In this thesis, the focus will be on endurance training.

Endurance is the ability to work at a relatively high intensity for a prolonged period of time (Skaset, 1976). This ability relies on the capacity to produce energy both anaerobically and, most importantly, aerobically. In the latter process, fats and carbohydrates are oxidised to CO_2 and water, releasing large amounts of energy. The lactic anaerobic process converts carbohydrates to lactic acid and produces energy at a fast rate, but with a restricted potential, as the accumulation of H^+ ions will interfere with the ability of muscles to contract (Fitts, 2008; Raymer, Marsh, Kowalchuk, & Thompson, 2004).

Endurance training aims to improve the fitness of the cardio-respiratory system, thus enabling the person to increase power output over prolonged time. The training effect is thought to be related to changes in both central factors (heart and lungs) and

peripheral factors (muscle fibre type, capillarisation, muscle enzymes and mitochondria).

This training can be either continuous or with intervals; in the latter type, the work intensity varies from high to low in a predetermined manner. Both forms of endurance training, continuous training (low intensity, high amount), and interval training (high intensity, low amount), are used successfully by athletes (Esteve-Lanao, Foster, Seiler, & Lucia, 2007). The split between high and low intensity is often determined by the lactate threshold, which is defined as the highest intensity that a person can endure over time without a concomitant increase in blood lactate concentration (Weltman, 1995). Training above the lactate threshold is often defined as high intensity training and that below the threshold as low intensity training. This is the definition used in this thesis. The factors that influence the training outcome are mainly the *intensity*, *amount* and *frequency.* Together they constitute the training *load*. In order to progress in aerobic fitness, intensity appears to be the most important factor (Wenger and Bell, 1986), unless large amounts of low intensity training are performed (Esteve-Lanao et al., 2007).

This training load is of a catabolic (degrading) type, which immediately reduces the performance level. Given suitable restoration, the anabolic process that follows training will increase the performance level not only to the previous state, but take it a bit higher, a process called supercompensation (Jakovlev, 1977).

Cardio-respiratory fitness is linked to physical health and physical activity. Morris and co-workers were the first to show the effect of physical activity on health in their “London Transport Society” investigation (Morris, Heady, Rafle, Roberts & Parks, 1953). Later, the Harvard Alumni Health Study by Ralph Paffenbarger, Jr. and

colleagues also demonstrated convincingly the positive effect of activity on physical health (Paffenbarger, Wing, & Hyde, 1978). From Norway, the HUNT study has produced good evidence for the positive health effect of active living (Tjønna, Lund Nilsen, Slørdahl, Vatten & Wisløff, 2010). People with an inactive lifestyle tend to have low MAP and more frequent cardio-vascular diseases than active people (Sesso, Paffenbarger, & Lee, 2000). From a health perspective, it will thus be of interest to enhance physical fitness in the population as a whole or in certain sub-groups.

Several investigations have shown that people with long-term substance abuse are in poor physical condition/fitness (Chalmers, Sulaiman, & Johnson, 1977; Gary & Guthrie, 1972; Frankel & Murphy, 1974; Murphy, Pagano, & Marlatt, 1986; Sinyor, Brown, Rostat, & Seraganian, 1982). This may partly be due to an inactive lifestyle and bad nutrition, but also because some substances, especially alcohol, have a deteriorating effect on heart and skeletal muscles, as well as on other vital organs of the body (Fernandez-Sola, Estruch, Grau, Pare, Rubin, & Urbano-Marquez, 1994; Fernandez-Sola, Preedy, Lang, et al., 2007; Kershaw & Guidot, 2008; Lundin, Hallgren, Landelius, Roxin, & Venge, 1986; Rubin, 1979;Urbano-Marquez, Estruch, Navarro-Lopez, Grau, & Rubin, 1989). Alcohol is high in energy content, so a high intake of alcohol can also lead to obesity. Substance abusers are neither much interested in doing physical training (Read, Brown, Marcus, et al., 2001).

Some investigations have reported estimates of MAP for substance-dependent people, revealing generally low values compared with an age-matched normal population. The most common test form is submaximal cycling, see Table 1.1.

Table 1.1 Physical fitness among substance-dependent people in treatment.

Authors	Year	Sex	Result	Work form (load)
Frankel and Murphy	1974	M	23 (test score)	Step test (submaximal)
Sinyor et al.	1982	M+F	29 ($ml \bullet kg^{-1} \bullet min^{-1}$)	cycle (submaximal)
Murphy et al.	1986	M	39 ($ml \bullet kg^{-1} \bullet min^{-1}$)	cycle (submaximal)
Sell and Christensen	1989	M+F	24 ($ml \bullet kg^{-1} \bullet min^{-1}$)	cycle (submaximal)
Collingwood et al.	1991	M	13:24 (min:s)	1-mile run (maximal)
Hellandsjø Bu	Unpub	M	29 ($ml \bullet kg^{-1} \bullet min^{-1}$)	cycle (submaximal)

M=males, F=Females. Tests which use a cycle ergometer as a test measure should have their results raised by 7% to be comparable with tests using running (Hermansen & Saltin, 1969).

The assessment of fitness shown in Table 1.1 is mostly of submaximal type, estimating instead of measuring aerobic power. Such methods do have a lower sensitivity and accuracy than results obtained from directly measured maximal tests (Segerström, Glans, Eriksson, Groop, Thorson, & Wollmer, 2008). There is therefore a need to establish the fitness level of substance-dependent people using direct measurements.

The lactate threshold is in many situations considered a better measure of endurance performance than MAP (Goodwin, Harris, Hernández, & Gladden, 2007), but up to now, no data on the lactate threshold of substance-dependent individuals has been published. This method of classifying fitness would be of interest, because the test is of submaximal character, and can thus be performed by a larger number of subjects. In our project the additional use of lactate threshold tests allows us to discriminate between training-induced changes in maximal and submaximal performance.

1.5 Improvement of aerobic power after training

How do substance-dependent people respond to training? Several studies have shown

that people dependent on alcohol respond to training in the same way as non-alcoholics (Palmer, Vacc, & Epstein, 1988). Table 1.2 summarises the results of six investigations with alcohol-dependent subjects. For substance-dependent people using other means of intoxication than alcohol, I have no knowledge of any empirical study concerning improvement in physical fitness following exercise. All previous studies have used simple test methods with a low level of precision, and I am not aware of any study where improvement in physical fitness has been measured precisely.

Table 1.2 Previously reported levels of improvement in aerobic power in persons abusing alcohol.

Author	Improvement (%)	Remarks: test type,duration,amount
Murphy (1970)	32	Step-test, 13 weeks, 7h/week
Frankel and Murphy (1974)	30	Step-test, 12 weeks, 5h/week
Sinyor, et al. (1982)	12	Å/R-test, 6 weeks, 5h/week
Murphy, Pagano, and Marlatt (1986)	12	Cycle-test, 8 weeks
Palmer, Vacc and Epstein (1988)	Not significant	Å/R-test, 4 weeks, 2.5h/week
Sell and Christensen (1989)	37	Å/R-test, 12 weeks, 1.7h/week
Collingwood, et al. (1991)	11	Mile-run, 9 weeks, 2h/week

h/week =hours pr. week, Å/R=Åstrand - Ryhming submaximal cycle ergometer test.

The improvements reported are not very different from improvement rates found in other training studies using a normal, untrained population. However, these rates of improvement are based on oxygen uptake estimations, not direct measurements of oxygen uptake.

1.6 The use of physical activity/training for mental distress

During the last decades, physical activity has been used as a method for treatment of various psychiatric disorders. The best documentation of a positive therapeutic effect is found for depression, and also to some extent for anxiety (Martinsen, 2008; Martinsen and Raglin, 2007). Physical activity has also been used in the treatment of substance abuse and dependence (Donaghy & Ussher, 2005; Stathopoulou & Powers,

2006). Most investigations in this field have methodological shortcomings, but the results still indicate that physical activity might be useful (Donaghy & Ussher, 2005). Clinical experience also points in the same direction.

In a review article on the use of physical activity for relieving mental distress, Stathophoulo and Powers (2006) stated that evidence exists for a beneficial effect of physical activity in alcohol- dependent subjects. Cross-sectional investigations show a strong inverse relationship between physical fitness and risk factors for abuse (Collingwood, Sunderlin, Reynolds, & Kohl, 2000). Ermalinski, Hanson, Lubin, Thornby and Nahormek (1977) found reduced craving for alcohol following physical training in a group with alcohol dependence. Review articles by Donaghy and Mutrie (1999), Tkachuk and Martin (1999), and Read and Brown (2003) point out that experienced well-being during and following physical activity can be an explanation for its usefulness in the treatment of dependence. Biddle and Mutrie (2007) claim that physical activity creates well-being sensations similar to those experienced in intoxication, and that this probably takes place through the production of β-endorphins. In addition, changes in lifestyle brought about by a more physically active life might also be of importance. These changes can also create more social support and lay the foundation for improved mental health through increased self-awareness and mastery. For those able to adhere to exercise, training seems beneficial for health over time. Murphy, Pagano and Marlatt (1986) found that a running programme significantly reduced alcohol consumption for heavy social drinkers, and that programme compliance was an important factor for success in the reduction of alcohol intake. Sinyor et al., (1982) found a reduced risk of relapse with physical training up to 18 months after programme termination. According to Donaghy and Ussher (2005),

there are many unanswered issues concerning the possible effect of physical training on substance dependence, and there is a great need for research-based knowledge.

Even though there have been several studies on the use of physical training in the treatment of substance dependence, these all have methodological weaknesses in that the measurements of fitness are not accurate, thereby making it difficult to investigate the relationship between fitness improvements and changes in mental distress.

1.7 Adherence and drop-out

The health gain from physical activity is dependent on regular activity over a long time, necessitating the training to become part of one's lifestyle. Unfortunately many patients do not manage to do this by themselves. They are able to perform regular physical activity while in treatment, but the major challenge is to keep up with the training after the rehabilitation is over. Drop-out is a serious problem in fitness programmes in general, such as cardiac rehabilitation programmes (Sanderson, Phillips, Gerald, DiLillo, & Bittner, 2003). Typically, 50% of the participants will have left the fitness programme within six months (Buckworth & Dishman, 2002). Massie and Shephard (1971) found that, in the general population, persons who dropped out of fitness programmes tended to be overweight, stronger, more likely to smoke and more extroverted compared with those who adhered to such programmes. Martin and Dubbert (1982) point to the importance of enthusiastic therapists, group-based workouts and low-intensity training for high treatment adherence.

There is a great need for creative methods for increasing exercise adherence in these patient groups. The use of "training partners" is one such new method, and will be evaluated in this thesis in terms of participant satisfaction. A training partner is a lay

person with some extra education in applied physiology and psychiatry, who trains together with the patient.

2 Aims of the dissertation

The first main aim of this thesis was to report, for the first time, fitness levels in substance- dependent persons based on maximal exertion tests and directly measured maximal aerobic power. In addition, we reported lactate threshold results from submaximal tests performed in an exercise mode adjusted to the training programme of the individual subject. **Research Question 1:** *What was the level of physical fitness in this group of substance-dependent people?*

The second main aim was to investigate the ability for improvement in fitness among substance-dependent individuals using measurements of the highest accuracy. With our test battery, we sought to distinguish between improvements of maximal and submaximal character. **Research Question 2:** *What was the degree of improvement in fitness among substance-dependent persons after a long and individualised training programme?*

The third aim was to investigate the relationship between the development of physical fitness and mental distress. For the first time, the fitness assessments were accurate enough to allow a closer inspection of the relationship between improvement of MAP and improvement of mental distress parameters, such as depression, anxiety, social phobia, general psychiatric problems and substance craving. **Research Question 3:** *What was the relationship between changes in physical fitness and mental distress in substance-dependent people?*

The fourth aim was to report how the use of training partners was received by the project participants. Having dedicated persons to train together with the patients is a new concept in dependence treatment, similar to having a personal trainer. **Research Question 4:** *How were the training partners perceived by the participants?*

3 Material and methods

The study can be described as a naturalistic, exploratory, prospective study, without a control group.

3.1 Participants

In all 55 participants were recruited to the project from December 2002 onwards. As seen in Table 3.1, the participants were recruited from a wide range of sources, but the majority were referred from the outpatient unit at the Department of Psychiatry, Sogn og Fjordane Central Hospital in Førde, Norway.

Table 3.1 Source of recruitment to the project (N=55).

Source	***Percent****
Psychiatric outpatient units	42
Recruitment source not given	35
Psychiatric institutions	16
Social security office	6
Local general practitioners	2

** Number rounding gives a sum that exceeds 100.*

Before baseline testing, eight persons had left the project. We have no data on these except age and sex, but they did not differ significantly from the remainder regarding these parameters. We have complete data on 47 persons at baseline (N 47). Of these, 36 completed the project, and we have a satisfactory data set from 33 (N 33) (70%), see Table 3.2.

Table 3.2 *Number of subjects in the different parts of the study, in total and as male/female ratio.*

	N	Males	Females	Per cent	Male/female ratio
Admitted	55	39	16		2.4
Drop-out pre baseline	8	5	3		1.7
Baseline	47	34	13	100	2.6
Drop-out	11	7	4	23	1.8
Insufficient data	3	1	2	6	0.5
Post-treatment	33	26	7	70	3.7

The characteristics of the participants at project start are shown in Table 3.3. The median age was 29 yr and the range was large at 33 yr. BMI was just below the usual limit for overweight, 25 $kg \bullet m^{-2}$. Most of the participants were out of work and 1/3 of them had had social security/rehabilitation benefit as their main income in the last month. About half of the group lived alone, without any partner. Nearly 50% of the participants had been in treatment for substance dependence before and they were all mixed users. Alcohol and cannabis were the two main primary intoxicants. A relatively small fraction of the group (~10 %) had used drugs or alcohol on a daily basis in the week prior to starting the project. One third reported that they had had a problematic use of alcohol/drugs for nine or more years. The prevalence of mental distress was substantial. At the admission interview, more than 80% reported to have had severe depression or anxiety previously, and about half admitted suicide. The differences between adherers and drop-outs were small and not statistically significant on most variables. However, drop-outs reported lower education level, less frequently received

previous treatment, more frequent use of cannabis and less often alcohol as primary intoxicant.

*Table 3.3 Description of the participants at baseline. N47 is the 47 subjects who started the intervention, drop-outs are the ones who left the project after baseline testing and N33 is those who completed the project. Results are mean (SD) or %. * denotes significant difference between drop-outs and the other groups. Data are obtained from the admission interview.*

	N47	Drop-outs (n=14)	N33	*P*=
Age (yr)	30 (10)	28 (9)	31 (10)	0.64
BMI ($kg \bullet m^{-2}$)	24.7 (4.6)	24.3 (5.1)	24.8 (4.4)	0.94
Compulsory school as highest education (%)	46	60*	41	**0.02**
Unemployed (%)	51	60	48	0.21
Rehab. benefit as main income (%)	28	30	27	0.89
Lives alone (%)	50	40	53	0.16
Not previously treated for dependence/abuse (%)	46	30*	52	**<0.01**
Alcohol as primary intoxicant (%)	43	20*	50	**<0.01**
Cannabis as primary intoxicant (%)	43	60*	37	**<0.01**
Daily use of intoxicants last week before treatment (%)	13	20	10	0.12
Problematic use ≥ 9 yr (%)	29	25	30	0.45
Suffered from depression previously in life (%)	88	90	87	0.80
Suffered from anxiety previously in life (%)	80	80	80	1.00
Attempted suicide previously in life (%)	49	44	50	0.66

3.2 The Førde Project

At the Department of Psychiatry at Sogn og Fjordane Central Hospital, Førde, a project was carried out from December 2002 to June 2005. During that period, all

patients admitted for in- or outpatient treatment for substance dependence were given the opportunity to take part in a systematic training project. Those who were interested were consecutively admitted to the project and assigned a training partner.

3.2.1 Training partners

To help the participants train regularly over time, and to enable them to conduct their own training afterwards, lay people were educated as training partners (Skrede et al. 2006). One training partner was assigned to each participant, and they took part in the activities together with their participant. An important goal was that the training partners should make themselves more or less superfluous during the project, enabling the participants to train by themselves. The training partners were given a course, with the objective of enhancing the competence of lay people in the local communities, so that they could take part in the project as training partners. The course had a total duration of 40 h, and the training partners paid a course fee of NOK 1300. The topics taught were:

- Substance dependence (8 h)
- Psychiatry (8 h)
- Applied physiology - practice (10 h)
- Applied physiology - theory (6 h)
- Anatomy and physiology (6 h)
- Communication (2 h)
- Home exam (2 h)

The Norwegian Confederation of Sport and The Olympic Committee approved the

course. The local social security offices hired the training partners and made them sign a declaration of silence. The training partners received the same financial support from the local authority as support persons in other projects not involving physical training (NOK 70 to 110 per hour). As of June 1st 2005, 286 approved training partners were registered in Sogn og Fjordane County, and 25 of the 26 local authorities could offer training partners. Eight of the project participants later took the training partner course themselves.

3.2.2 My involvement in the project

The project was planned and led by Atle Skrede and Harald Munkvold from the Department of Psychiatry, Sogn og Fjordane Central Hospital, Førde. My task was to give the participants physiological tests at admittance and discharge.

3.3 Data collection

3.3.1 Assessment of mental distress

The level of mental distress was assessed by six inventories, which were in regular use at the Department of Psychiatry.

- **The Beck Depression Inventory II** (BDI-2) has 21 items, which are scored on a 0–3 scale (Beck, Steer, & Brown, 1996). It is a self-report measure used for assessing level of depression, and the sum score ranges between 0–63. Higher scores indicate more severe depressive symptoms. The cut-off scores normally used are: 0–9 normal range, 10–19 mild depression, 20–29 moderate depression, 30 and above severe depression.

- **The Beck Anxiety Inventory (**BAI) has 21 items scored on a 0–3 scale, with a score range of 0-63 (Beck & Steer, 1993). Normally used cut-offs are: 0–7 minimal level of anxiety, 8–15 mild anxiety, 16–25 moderate anxiety, 26–63 severe anxiety.
- **The Brief Social Phobia Scale** (BSPS) is an 11-item measure of social phobia symptoms (Davidson, Potts, Richichi, Ford, Krishnan, Smith, & Wilson, 1991). The severity of each item is rated on a 0-4 scale, and a total score and three subscores are usually calculated: Fear, avoidance and physiological reactions. The total score range is 0–72, scores above 20 usually indicating a need for treatment.
- **The Symptom Check List-90 Revised** (SCL-90R) (Derogatis, 1983) is used to assess general mental distress, and 90 items are rated on a 0–4 scale. The mean score of all items is called the Global Severity Index (GSI). GSI scores above 0.7 are often considered to be outside the normal range.
- **The Drug Abuse Screening Test 20** (DAST-20) (Skinner, 1982), is a 20-item self-administered inventory. Questions are scored as yes = 1 and no = 0. Screening test scoring ranges are: 0 none reported, 1–5 low level, 6–10 moderate level, 11–15 substantial level, 16–20 severe level.
- **The Short Michigan Alcoholism Screening Test** (SMAST) (Selzer, Vinokur, & Rooijen, 1975; Pokorny, Miller & Kaplan, 1972) is a 13-item inventory with yes/no response alternatives. Three or more “yes” responses indicate alcoholism.

In addition:

- **Attitudes to alcohol/drug use** from Beck, Wright, Newman and Liese (1993) comprises 20 statements on attitudes to alcohol/drug use. This questionnaire was translated into Norwegian by Alice Mais at Førde Hospital. This was not an authoritative translation with a translation/back-translation procedure. The inventory was administered to the participants at start and discharge. The patients scored each statement on a 1–7 scale, where 1 was total disagreement, 7 was total agreement and 4 was neutral. The statements have been grouped into four categories in this thesis: "Positive attitude to alcohol/drug use" (statements 1, 2, 8, 9 and 20), "Alcohol is a necessity" (statements 3–7, 11, 12, 14-17 and 19), "I'm an unworthy person" (statements 10 and 18), and "Alcohol/drug use is unproblematic" (statement 13).

3.3.2 Physical training

Physical activity was recorded by the training partner in a training log, where duration, type of training, intensity and special remarks were noted. In some training sessions heart rate was also recorded with a Polar heart rate monitor (S610i) and Polar Precision Performance software (Polar Electro OY, Kempele, Finland).

3.3.3 Evaluation of the training partner

The training partner was evaluated after the project on the basis of two questions with a score ranging from 1 (not satisfied/not important) to 4 (very satisfied/very important).

3.3.4 Rating of perceived exertion

For rating of perceived exertion, the 15-point Borg scale (Borg, 1970), which is based

on the 21-point scale by Borg (1962) was used. This scale uses category scaling in a linear fashion with semantic anchoring for every second score. The score of 6 is "nothing" and 20 is "maximal exertion". Before start, the participants were instructed on how to use the scale with a standardised explanation and additional extra information if needed. The scale has been extensively validated (Borg, 1982; Noble & Robertson, 1996).

3.3.5 Physiological measurements

Protocols

Lactate profile tests were conducted on either an ergometer cycle or a treadmill. A cycle ergometer was chosen for participants who had stated that cycling would be their preferred type of training. It is important that testing mode and training mode match, as the heart rate at threshold load (HR_{LT}) is activity-specific. Using a different test modality from the primary training form would have made the threshold testing concept less valid, due to the specificity of training. The cycle test consisted of five-minute work periods with no rest in between. The starting load was 1 to 2 W kg^{-1} body mass and the cadence was 75±2 RPM. The increase in step load was 30 W every fifth minute, without a pause for blood collection. The test continued until the blood lactate concentration (BLC) was about 4 mmol/L^{-1}. On the treadmill, the first step was of 10 minutes duration, at a speed approximately equal to 65% of HR_{max} (1.67 to 2.22 m s^{-1}). Subsequent steps lasted five minutes and the speed increased by 0.042 m s^{-1} each time (1.5 km h^{-1}). Between the steps, a 30 s rest period was used to collect blood. The test was terminated when the BLC was close to 4 mmol/L^{-1}.

For both the cycle and treadmill testing, the lactate threshold was calculated using polynomial regression and a threshold point as baseline BLC value + 1.5

mmol/L^{-1} (Helgerud, Ingjer, & Strømme, 1990). After a 15-25 min resting period, the Maximal Aerobic Power test started. This test was performed on the treadmill according to the Bruce protocol (McArdle, Katch, & Katch, 2010, p. 238) until voluntary exhaustion. The objective criteria for exhaustion included a respiratory exchange ratio (RER) >1.00 or a levelling off in the oxygen uptake despite an increase in load. The subjective criterion was an observed inability to continue the test. Two of the criteria had to be fulfilled to have a valid maximal test.

Equipment

We used a Metamax I metabolic analyser (Cortex Biophysik, Leipzig, Germany) to record oxygen uptake and CO_2 production. Data was recorded every tenth s and averaged over 30 s. The treadmill was a Woodway PPS 55 (Woodway GmbH, Weil am Rhein, Germany). The cycle ergometer was a Monark 824 E (Monark Exercise AB, Vansbro, Sweden). Heart rate was recorded with Polar heart rate monitors (Polar Electro OY, Kempele, Finland). Blood lactate concentration was measured with a Lactate Pro LT-1710 analyser (Arkray Inc, Kyoto, Japan).

3.3.6 Responders and non-responders

In papers II and III below, the participants were divided into responders and non-responders, according to how much the MAP had changed. If the change in aerobic power was less than 150 ml•min^{-1}, they were considered non-responders. This was done to avoid spurious improvements from random analyser variation and human day-to-day variance (Kuipers, 1983).

3.4 Ethical considerations

The participation was voluntary. All participants signed an informed consent, stating that they could leave the project at any time without giving any reasons for doing so,

and without this having any negative consequences for future treatment. The project was an integrated part of the clinical activity at the institutions concerned, and was initially a quality improvement project, aiming at monitoring and improving the treatment at the particular institution. For this reason, it was not considered necessary to apply to the Regional Committee for Medical and Health Research Ethics for approval. However, the committee was consulted, and their responses were taken into consideration in the final project outline. The whole project was conducted according to the Declaration of Helsinki on medical research on humans (http://www.cirp.org /library/ethics/helsinki/).

3.5 Statistics

Subjects with missing data from pre- to post-treatment (n=3) were excluded from the project. Data in this thesis are expressed as mean (SD) unless otherwise specified. Effect sizes are expressed as Cohen's d (Cohen, 1988) and 95% confidence interval (95% CI). For nominal data, mode is used as the measure of central tendency. Parametric data were compared with a t-test or ANOVA. Data that failed the Kolmogorov-Smirnov procedure (Lilliefors correction) for normality were analysed with the Wilcoxon-Mann-Whitney U-Test or ANOVA on ranks. Fischer's exact test and chi-square were used for comparison of proportions. The level of significance was set at $p \leq 0.05$. Statistical software used included SPSS v.14-17 (SPSS Inc., Chicago, USA), Sigma Plot v. 10 (Systat Software GmbH, Erkrath, Germany), Winks SDA v. 6.0.8 (TexaSoft, Cedar Hill, TX, USA), the CEM effect size calculator, Centre for Evaluation and Monitoring, Durham University, United Kingdom: http://www. cemcentre.org/renderpage.asp? linkID=30325017) and Quick Calcs, Graph Pad Software, La Jolla, CA, USA: http://www.graphpad.com/quickcalcs/ index.cfm.

4 Results

The maximal aerobic power was 37 ml•kg^{-1}•min^{-1} at project start, which is subnormal. The improvement in MAP was modest at 5%. Threshold performance improved more, by 11%. All scores of mental distress were significantly reduced during the project. Results of screening tests for alcohol and substance dependence did not change during the project. Those who responded to the training by increasing their MAP with more than 150 ml•min^{-1} experiences larger reductions in mental distress. Finally, the use of training partners was positively evaluated by the patients.

4.1 Research Question 1

The MAP was 2.78 L•min^{-1} or 37 ml•kg^{-1}•min^{-1} at baseline. The maximal heart rate at 188 was close to the 220-age rule. The participants were able to utilise 74% of their MAP before blood lactate started to accumulate. The cycling power at lactate threshold was 86 W and threshold running speed was 2.05 m•s^{-1}. RPE at blood lactate threshold was 13, i.e. “somewhat hard”.

Some of the results for cycling and running differed statistically: Cycling HR_{LT} was significantly lower than running HR_{LT} (125 (14) vs. 144 (15), p=0.02).

4.2 Research Question 2

After 36 (16) weeks of training time (time between the two testings), the post-treatment testing did show changes in several of the measured parameters. The group had a mean project time (time between inclusion and discharge) of 11 (2) months and in that time span they performed 301 (94) h of training, of which 26 (15) % was high intensity training. The shortest time in the project was 8 months and the longest 12 months. The amount of training ranged from 150 h to 500 h. Some of the participants had no training time above threshold heart rate, the maximum being 60 %.

4.2.1 Development of maximal aerobic power

The improvement in MAP was statistically significant for both ways of expressing it (absolute and relative to body mass). The improvement was 5% for MAP expressed in L• min^{-1}, and 4% for MAP relative to body mass ($p<0.02$).

4.2.2 Lactate threshold performance

During the training period the subjects improved in performance at lactate threshold. Their cycling power increased by 16%, and their running speed by 8%. The combined improvement for the group was 11%. Their heart rate at lactate threshold (HR_{LT}) increased significantly whereas their rating of exertion did not change notably. The rate of improvement in MAP and threshold performance correlated strongly. The only correlation that did not prove significant was the change in MAP relative to body mass vs. change in cycle power.

4.3 Research Question 3

All main indicators of mental distress showed significant changes at post-treatment testing. Eight participants moved into the normal range score for depression and anxiety at post-treatment. With regard to social phobia (BSPS), the reduction was more modest; three subjects moved from above to below the cut-off for treatment during the project, but the reductions in social phobia were significant for the total score and the sub-scores of avoidance and physiology. Regarding the SCL90-R-GSI score, at post-treatment, the number of subjects who scored ≥0.70 was reduced by five, a highly significant reduction and also the largest in percentage terms.

The change in general psychiatric problems (GSI) was significantly related to the reduction of BDI (r=0.57, p<0.01), BAI (r=0.66, p<0.01) and total BPSP (r=0.53, p<0.01). Total BSPS change also correlated significantly with change in BAI (r=0.63,

$p<0.01$) and with BDI ($r=0.35$, $p=0.05$).

No statistically significant reduction was observed in the mean score of alcohol abuse recorded with the SMAST inventory. Actually, at post-treatment time, one more person scored ≥ 3 on the SMAST test. In DAST-20, the number of persons who scored ≥11 was reduced by three at post-treatment.

The responders experienced significant reductions in most of the measures: BDI, BAI, BSPS total, fear and avoidance and GSI ($p\leq0.02$ for all). On the other hand, the non-responders only improved significantly with respect to the GSI ($p=0.02$).

4.4 Research Question 4

At discharge the participants were asked about how they had experienced the project and the training partners. In all, the participants were satisfied with the project and also gave the training partners a high score. The responders were more positive to the project than the non-responders.

4.5 Synopsis of articles

Paper I

Mamen, A. & Martinsen, E.W. (2009). The aerobic fitness of substance abusers voluntary participating in a rehabilitation project. *Journal of Sports Medicine and Physical Fitness*, *49*, 187-93.

The aim of the study was to assess the physical fitness of a group of substance dependent persons using direct, maximal testing and testing of blood lactate threshold. Forty-seven patients with substance dependence (34 males and 13 females) taking part in a rehabilitation project served as subjects. Mean age and SD were 30.5±9.2 yr for the males and 29.7±10.2 for the females. They performed a lactate threshold test and a

maximal aerobic power test (MAP). The threshold test was performed on either a cycle ergometer or a treadmill, and the MAP test followed the Bruce protocol. The range in aerobic fitness was large, and the mean result of MAP was somewhat higher than previously estimated in groups of alcohol dependent patients at similar age: 39±10 and 31±8 ml•kg^{-}1•min^{-1} for males and females respectively. The mean heart rate at blood lactate threshold did not differ between the sexes and was 74±7% of maximal heart rate (HR_{max}). There were no statistically significant differences between the users of different intoxicants on either MAP or lactate threshold. **Conclusion**: This group of substance dependent persons showed a wide range in physical fitness. The mean MAP was better than previously assessed in groups of substance-dependent people, using indirect methods. The primary intoxicant had no significant influence on the fitness.

Paper II

Mamen, A. & Martinsen, E.W. (2010). Development of aerobic fitness of individuals with substance abuse/dependence following long-term individual physical training. *European Journal of Sport Science, 10,* 255-262.

Thirty-three persons with substance dependence who completed a training programme of 11 (2) months duration were evaluated on development of aerobic power and performance on a lactate profile test using directly measured MAP and measurement of blood lactate concentration. The group improved moderately on MAP, 4[2]%, mean[SEM], (p=0.020), and on lactate threshold performance, with 7[3]% for running (p=0.03) and 13[5]% for cycling (p=0.03). The low improvement rate compared to other studies may be attributed to the use of directly measured MAP, as indirect tests used in previous studies are prone to overestimation of the improvement. Those who

improved their aerobic fitness did more of the training above threshold heart rate than those who did not improve (p=0.006), although the amount of training during the project, 315 (97) h vs. 279 (89) h (p=0.30), did not differ significantly. **Conclusion:** The project is the first to show that direct testing of fitness using maximal tests is feasible for this patient group. The project is also pioneering in the use of training partners to help patients with substance abuse and dependence to perform large amounts of training.

Paper III

Mamen, A., Pallesen, S., & Martinsen, E.W. (in press).Changes in mental distress following individualised physical training in patients suffering from substance dependence. *European Journal of Sport Science.*

Thirty-three persons with substance dependence and psychiatric disorders underwent a high load physical training programme with the help of dedicated training partners who motivated them to exercise and controlled the intensity of the workouts. To monitor fitness, direct measures of MAP and lactate threshold performance were assessed before start and at the end of the project. An average of 300 h of training was performed during the project. The participants improved moderately in physical fitness, by 4-5%. Levels of depression, anxiety, social phobia and general psychiatric problems declined significantly during the project. The results of screening tests for alcoholism and drug abuse did not change significantly. Due to the use of accurate physical tests, it was possible to objectively distinguish between responders who improved in physical fitness and non-responders. **Conclusion:** The responders reported greater improvements in mental health at discharge than the non-responders.

5 Discussion

5.1 Physical fitness in substance dependent people with additional mental distress

Statistically significant changes were seen in the physical fitness of the participants during the period they stayed in the project.

5.1.1 Maximal aerobic power

The physical fitness of the group, expressed as MAP, revealed that the patients did not match the MAP level of healthy age-matched peers. The males could be classified as being at the top of the "below average" classification, or the 40th percentile, whereas the females came at the bottom of the same classification, belonging to the 30th percentile using the YMCA classification (Golding, Myers & Sinning, 1989). This is somewhat better than the findings of previous studies on alcohol abusers. One reason for this might be the more accurate method we have used to assess physical fitness. In addition, substance abusers probably are a very diverse group, showing a large range in fitness, and it is known that earlier fitness level, abuse duration and type of intoxicants may influence the fitness.

This heterogeneousness with regard to fitness was also apparent in our group, and the least fit subjects would be classified as "poor" in the YMCA classification. The fittest individuals would be classified as "average" to "high" and they would not find even strenuous work too demanding (Jørgensen, 1985). The fitness level was not related to the duration of abuse or to the type of intoxicants used.

5.1.2 Lactate threshold performance

Testing substance-dependent people with a blood lactate threshold test is a novel

procedure. There is thus a lack of comparable data. At the baseline cycling test the mean threshold power was modest, with a large range in performance, 159 W. The same can be said of the running performance range, 1.13 $m \bullet s^{-1}$. The lowest threshold cycling power was 50 W. Such a load requires an oxygen uptake of about 1.0 $L \bullet min^{-1}$ or 3.5 METs, which is usually labelled as low to moderate work intensity (Åstrand, Rodahl, Dahl & Strømme, 2003, p.282). The lowest running speed, 1.36 $m \bullet s^{-1}$, would require about the same oxygen uptake and MET value.

The 25th percentile of utilisation was 69%. The blood lactate threshold sets a limit for how much of the maximal aerobic power can be utilised for a prolonged time. According to McArdle, Katch and Katch (2010, p. 163), the average utilisation for untrained persons is between 50 and 60% of MAP, which corresponds to 66% to 73% of HR_{max}. The least fit quartile of our subjects would thus not be able to sustain moderate to hard work for a long time. The fittest quartile demonstrated utilisation of 78%, which is found in trained recreational endurance athletes, and would allow a person to sustain hard work for a long time. Thus, there is a wide range in ability to perform physically demanding work in our group. The females in this investigation managed a mean utilisation of 73%, and the males 71%, at baseline. These values are within the upper range of the untrained population (ibid.).

5.1.3 Rating of perceived exertion

Rating of perceived exertion is a common way to grade exercise (Borg, 1998) and many of the factors influencing the rating have been investigated. Even though lactate threshold is the subject of both criticism and a variety of definitions (Weltman, 1995; Faude, Kindermann, & Meyer, 2009), much research has focused on RPE at blood lactate threshold load (Weltman, 1995). The mean rating we find here at baseline, 13

(2), is in the lower range of that found in the general population, where ratings of 13–15 are common (Noble and Robertson, 1996; Purvis & Cureton, 1981). It might be that our group, with a long history of substance abuse, have experienced hard life conditions, and have therefore become "toughened", so their rating of physical exertion is low. In the same group we have previously shown that those with additional ADHD (nine subjects), rated the effort at threshold point significantly lower than those without ADHD (Mamen, Skrede, Munkvold, Martinsen, 2009). The substantial number of persons with additional ADHD might therefore have contributed to the low threshold RPE.

5.2 Development in physical fitness following training

During the project, an unusually high number of training hours was performed. The mean value was 300 h, and one participant had logged 500 h of training in less than a year. Elite athletes in endurance events regularly train 1000 h or more in a year, but for recreational athletes, the amount is considerably less; 200 to 600 h would be an educated guess. In the light of this, the number of training hours performed in the project is exceptional, and much higher than in any of the other studies that have reported the training amount in programmes for substance abusers (Murphy, 1970; Murphy et al., 1986; Frankel & Murphy, 1974; Sell & Christensen, 1989; Palmer et al., 1988).

The training was recorded by the training partner and also supplemented with additional heart rate recordings in some of the sessions. A major goal for the project was to individualise and optimise the training intensity through the use of accurate testing. The blood lactate threshold test was used to split the training into low and high intensity, and with the help of Polar heart rate monitors, the intensity could be

controlled. Analyses of the electronic training diary (Precision Performance, Polar Electronic OY) indicated that about 25% of the training was of high intensity. This fits well with current opinions on how to structure endurance training (Fiskerstrand & Seiler, 2004; Seiler & Kjerland, 2006; Esteve-Lanao et al., 2007; Midgley, McNaughton, & Wilkinson, 2006).

There was a considerable range in intensity. Some individuals did all of their training below threshold heart rate, while others had 60% of the training above this intensity. The debate about the efficacy of high intensity training has been going on for several years. Research has shown that the stroke volume may not plateau in some individuals in certain forms of exercise (Rowland, 2009). This indicates that training at a high intensity, 90-100% of MAP, is most beneficial in developing MAP, at least in some individuals and work forms. In a review article, Butcher and Jones (2006) found evidence for using high intensity interval training in COPD patients, but Midgley, McNaughton and Wilkinson (2006) could not find evidence for the supremacy of high intensity training in developing MAP. In well-trained endurance athletes, Esteve-Lanao et al. (2007) found that an approach with most of the training in the low intensity zone (80%), some training in a medium zone (12%), and some in the high intensity zone (8%) gave better improvement in a 10 km run test than training with more medium zone intensity (25%), equal high intensity zone training (8%) and reduced low intensity zone training (67%). The classification of training intensity was based on ventilatory threshold (VT) measurements. Hansen, Blanchard, Rodgers and Bell (2003), using untrained persons, also structured a 12-week training programme based on VT measurements. (VT = the point during graded exercise in which ventilation increases disproportionately. Usually two thresholds can be found, the first

(VT1) is where EqO2 (V_E/VO_2) increases while $EqCO_2$ ($\dot{V}E/\dot{V}CO_2$) is constant. The second (VT2) is where $\dot{V}CO_2$ increases non-linearly in relation to $\dot{V}O_2$). Equal improvement in MAP was found for training at both VT1 and VT2, showing that the use of threshold training is also beneficial for untrained persons. VT1 is taken to correspond roughly to the lactate threshold (Weltman, 1995), thereby indicating that the use of lactate measurements is also viable for setting the training intensity.

5.2.1 Change in maximal aerobic power

MAP did increase modestly after a long training period with a high load of training, but almost half of the participants did not experience any improvement (n=14). Lack of development in fitness after a training project can be caused by a number of factors. As Wenger and Bell (1986) pointed out, training intensity is important, and the responders in our group did train with a higher intensity than the non-responders. This finding is also in line with the findings of Wisløff, Ellingsen, and Kemi (2009).

When a high load of training has been performed and progress is less than expected, overreaching or overtraining should be considered. The basis for such an idea is that the organism reacts to training with catabolism (tearing down tissues) and uses the recovery period for an anabolic phase (building up tissues), which subsequently leads to a super-compensation above the previous level (Jakovlev, 1977). If this recovery period is not satisfactory in duration or quality (nutrition, sleep, mood disturbances, etc.), the super-compensation stage is not reached, and a continuous catabolic progress may be the result. This may lead to overreaching or subsequently to overtraining (Halsom & Jeukendrup, 2004).

A high amount of training is in no way a guarantee of success, as the training load also has to be increased during the training process (De Lorme and Watkins,

1948). A lack of increase in training intensity will often result in the training not giving the expected results (Wenger & Bell, 1986). So, the lack of progress may be due to inadequate restoration or to an inadequate training stimulus. Since we have heart rate data from many of the training sessions, and the sessions were guided by a training partner, thereby giving an adequate training stimulus, it is most reasonable to claim that there might have been overreaching by at least some of the participants.

In training studies it is not unusual to find a rapid increase in MAP, which then plateaus, or even slightly declines, as the training goes on. Since we have used end-point measures, we do not know what has happened in between the tests.

The progress in fitness found here could probably have been attained in a shorter time. Whether this means that the project should have been shortened is another matter. Other factors, such as building good training habits, may take a longer time than just increasing MAP.

5.2.2 Lactate threshold/utilisation

The performance at lactate threshold increased significantly, and the subjects were able to utilise a larger part of their MAP before lactate started to accumulate at post-treatment. The range increased from baseline to post-treatment, illustrating the uneven development in the group. As for the level of performance, the least fit would still be unable to do hard physical work (25^{th} percentile: 82 W and 2.00 $m \bullet s^{-1}$), whereas the fittest would find such work highly tolerable (75^{th} percentile: 136 W and 2.53 $m \bullet s^{-1}$).

5.2.3 Rating of perceived exertion

The RPE score did not change much from baseline to post-treatment testing, even though the heart rate and work load increased significantly. The mean (SD) RPE at

post treatment was 13 (2) which equals “somewhat hard”. This shows that the use of RPE can be very useful in fitness training, as the rating at threshold is constant throughout the training period. Foster, Fitzgerald and Spatz (1999), found the RPE to be stable at aerobic and anaerobic threshold from spring to autumn in speed skaters, even though the performance improved significantly. By having the participants exercise at a given RPE, the load will increase as fitness improves.

5.3 Relationship between changes in fitness and mental distress

There was a considerable change in the level of mental distress, and the size of the change mimicked the change in physical fitness.

5.3.1 Responders and non-responders

Responders and non-responders did not differ with regard to mental distress at baseline. They both remained in the project for an equal amount of time and performed a similar amount of training. Both the groups experienced reductions in mental distress, but the responders showed greater reductions on most of the measures. This is contrary to the findings of Sexton, Mære & Dahl (1989) in their investigation of the effect of low and high intensity training on depression, anxiety and general psychiatric problems in a group of neurotics. In their study, gains in aerobic power did not correlate with improvements in mental health at study end; however, those with the greatest gain in MAP showed a reduced anxiety level six months later. They also speculated that initial high intensity training could inhibit future exercise activity, since the hard training could be experienced as negative. However, their total training load and intensity were much less than in our study.

At baseline, the responders in our study had significantly lower BMI than the non-responders. Being lighter may be an advantage in endurance training, especially if weight-bearing activities are undertaken. Thus, the responders had an advantage in this type of training, and this may provide a better feeling of mastery. This is an interesting finding, and future interventions with physical activity should also pay attention to body mass/BMI when planning activities.

In training programmes, the individual response to the training usually varies. Some improve more than others. This can be related to a number of reasons; adherence to the protocol, starting level of fitness and genetic factors such as muscle fibre composition. In one recent study, Vollaard and colleagues (Vollaard, Constantin-Teodosiu, Fredriksson, et al., 2009) investigated the adaptations in maximal aerobic power in 24 sedentary males during 24 training sessions over six weeks. The rate of improvement ranged from -2% to more than 30%, yet the "low responders" on MAP were not consistently low on other parameters measured. Using RNA signatures, Timmons, Knudsen, Rankinen, et al. (2010) were able to explain nearly 50% of the response to endurance training. However, Lucía, Morán, He and Ruiz (2010), warn against placing too much weight on genetics when predicting athletic performances and responses to training.

Stalling, (overreaching) or declining (overtraining) in performance level may occur due to an unfavourable relationship between training load and restoration; although the non-responders trained at a lower intensity than the responders, their training load may have been too high relative to their restoration, which includes nutrition, sleep and other loads from everyday living. If they did not successfully recover back to previous levels before a new training load started, they would

eventually perform worse.

The responders also reported a greater improvement in physical and mental health, showing that their personal experience was in line with the objective physiological measurements. The responders had also improved their relations with family and friends more than the non-responders, and scored higher on a question on improved quality of life at the discharge interview.

The responders participated more frequently in group training sessions, which Martin and Dubbert (1982) pointed out to be an important factor for positive treatment outcome. When asked about how much they were going to exercise by themselves after the project had ended, the responders answered significantly more positively than the non-responders. Such a difference is indicative of motivational differences between the two groups for continued physical training, which may be based on differences in experience of mastery and self-efficacy during the project. The responders showed a large effect size (0.73) on change on the total BSPS score, indicating a reduction in symptoms of social phobia at post-treatment, whereas the non-responders had a low effect size (0.12), signifying no substantial change in this measure. Why the social phobia symptoms were so persistent among the non-responders we cannot say, but the above-mentioned difference in body mass may have made the non-responders more vulnerable to social pressure and their perceived expectations from others.

5.3.2 Alcohol and drug dependence

The SMAST and DAST-20 scores did not change significantly from baseline to post-treatment, indicating that the dependence on substances did not improve through the training programme. Therefore, our data are contrary to the report by Sinyor et al. (1982), who found an enhanced abstinence among those who had undergone physical

training. One reason may be that during a long intervention period, as employed in the present study, there is a high chance of relapse, irrespective of the training programme. One must also be critical to the tests used to measure substance dependence, since they are all based on self-report and they are not shown to be sensitive to change. The use of other tests, such as the AUDIT (Babor, Higgins-Biddle, Saunders & Monteiro, 2001; Cassidy, Schmitz, & Malla, 2008) and the DUDIT (Berman, Bergman, Palmstierna, & Schlyter, 2005), might have been more appropriate, as they can be more sensitive to change.

5.4 Evaluation of the training partners

The relationship between the training partner and the participant was important for a positive development during the project. The partners had undergone special training to deal with dependent patients and had the task of transmitting their knowledge of applied physiology to their project participant. The training partners were highly valued by the participants. It can therefore be argued that such a “personal trainer” may be a relatively inexpensive way to promote participation in, and adherence to, a long-term training programme. The social part of the project was likewise highly valued. Participation in group training sessions was high, again pointing to the importance of social factors in dependence rehabilitation projects, as mentioned by Martin and Dubber (1982).

5.5 Methodological considerations

The choice of methods is important for the ability to answer research questions. To make assumptions of causality, longitudinal studies have to be performed, and a control group must be present.

5.5.1 Study design

This project can be looked upon as a naturalistic, prospective and exploratory study. Such a design is appropriate for investigating changes over time, and how different variables change in relation to each other. This design is more suitable for generating than testing hypotheses. There is a low current level of evidence-based knowledge for treating substance-dependent people with physical training, which does not enable us to make firm decisions on how to conduct the training and on whom. The findings here bring new knowledge to the field, and allow for a more efficient use of training in this patient group.

5.5.2 Validity and reliability

The validity of a measurement is how exactly the measure matches the real object of interest: Do you measure what you want to measure? Reliability is how faithful the measure is: Does it measure the same thing each time under the same conditions? High validity and reliability are of the utmost importance for a researcher, as they are prerequisites for making firm conclusions.

It is usual to divide validity into internal and external validity. The first is threatened by all possible alternative explanations for the outcome, and external validity is the ability to generalise the findings beyond the participants and the study settings. Usually there will be a inverse relationship between the two: A high internal

validity, caused by a rigorous laboratory setting, may give a low external validity due to the very special setting of the experiment, and if the external validity is very high, as when the subjects have been "free-ranging", then you will generally find other possible explanations for your results, resulting in low internal validity.

The most important threats to validity are the effects of subject history, subject maturation, instrument decay, regression to the mean, subject drop-out (experimental mortality) and the influence of the testing experience itself on the final results.

Subject history

Any incident that happens to a participant during an experiment may influence the outcome. We did not record such events, but it is conceivable that if a participant received positive comments from friends and family, this may have led to more energy to pursue the project goals, and thus a better result. Conversely, negative experiences such as drug use, not necessarily directly connected to the project, or other negative life events, may lead to a poorer final result.

Subject maturation

People change with time. This is mostly evident in children, but adults too change with the passing of time. This project lasted almost a year, which is a long time for an experiment. It is therefore quite plausible that our participants also experienced some maturation, which may have affected the results.

Instrument decay

Instruments may change due to ageing, especially in long lasting experiments/studies. However, in measurements of physiological parameters, this detrimental effect is non-existent, both due to regular calibration procedures and the use of zirconium electrodes for measuring O_2, and the infra-red technique for CO_2 measurements. Mental distress

was recorded by self-report, and the participants may have experienced some changes that made them evaluate the questions differently at baseline and post-treatment testing.

Regression to the mean

When groups are based on extreme scores (best/worst), there is likely to be a tendency to move towards the mean for both groups, i.e. if measured a second time, the best group will perform worse, and the worst group will improve. In our setting, this effect may be inflating the difference between the responders and the non-responders. A second measurement might have resulted in a smaller difference. It is also evident that since the group as a whole was in poor shape at baseline, regression towards the mean would imply some improvement at post-treatment. The regression to the mean effect can be minimised by using a randomised control group, and using analysis of co-variance when there are group differences at baseline.

Experimental mortality (subject drop-out)

The problem of bias due to subject drop-out is a real threat to all experiments, especially experiments/studies lasting for a considerable amount of time. This will be dealt with in the next chapter (5.5.3).

The effect of testing

Being an object of scrutiny, being tested, may have an impact on how the subjects react. The classic "Hawthorn experiments" by Enrice Mayo and co-workers in the late 1920s and early 1930s showed that receiving attention from scientists made the workers feel special, which in turn affected their production. However, Jones (1992) has questioned the whole effect.

Validity in psychology

Cronbach and Meehl (1955), in a "classic" paper, divided validity into four categories: Construct validity, content validity, concurrent validity and predictive validity.

Construct validity

The construct validity of an instrument is how well the outcome reflects the intended object of measure. In our case, how well does a MAP test reflect the physical fitness of a subject and how well do lactate threshold results reflect fitness?

Content validity

The content validity refers to how well a test/instrument covers all aspects of the construct. Is fitness adequately defined by the tests chosen, and do the mental distress inventories cover all aspects of the distress they are supposed to measure?

Concurrent and predictive validity

These two validity types refer to how well a test score and a criterion score correlate. Concurrent validity concerns the here and now correlation between the scores on the instrument and an external criterion, whereas predictive validity refers to the relationship between scores on the instruments and future measures of the criterion. For us, it could be how well a MAP result correlated with e.g. 3000 m running performance, either immediately, or at some later time (given that MAP could be kept at its present level).

Validity of measurements of physical fitness

- MAP

It has long been textbook knowledge that oxygen uptake is linearly related to heart rate, work load and ATP production (Åstrand et al. 2003). The metabolic analyser has been validated and compared to other lactate analysers with good results (Medbø, Mamen, Holst Olsen, & Evertsen, 2000; McLean, Norris, & Smith, 2004; Mukherjee & Chia, 2006).

- Rating of perceived exertion

Borg´s RPE scale has been in use since 1970, and is the most common scale for assessment of physical exertion. The scale has been subjected to numerous validation and reliability studies (Noble & Robertson, 1996). Chen and co-workers (2002) conducted a meta-analysis on more than 60 journal articles in English on the validity of RPE and found the validity coefficients (Pearson product-moment correlation) for RPE and six physiological parameters (heart rate, blood lactate concentration, % $\dot{V}O_{2max}$, $\dot{V}O_{2max}$, $\dot{V}_E$ and respiration rate) to be from 0.57 to 0.72 (Chen, Fan & Moe, 2002).

Validity and reliability of mental distress measurements

The mental distress inventories, BDI, BAI, SCL-90-R, BSPS, and the substance abuse tests DAST-20 and SMAST were all used as self-report inventories at inclusion and discharge. This can lead to user manipulation in that the respondents may enhance or diminish responses to certain topics. Both their own expectations and expectations from others can have an impact on how the inventories are completed, thereby creating a bias and reducing their validity. In the context of our study, we find Goffman's theory of self-presentation interesting (Goffman, 1958).

- BDI

The Beck Depression Inventory has a long history of use, and Richter, Werner, Heerlein, Kraus, and Sauer (1998) found the BDI to have a high validity. Likewise, Hesse (2006) found the BDI to be stable and of predictive validity on a sample of opiate-dependent patients. Very severe drug use could, however, reduce the stability of the BDI.

- BAI

The Beck Anxiety Inventory has been validated in several settings, with a wide span of patient groups, and found to give a valid measure of anxiety (Beck & Steer, 1990; Enns, Cox, Parker, & Guertin, 1998; Gillis, Haaga, & Ford, 1995; Hewitt & Norton, 1993; Lykke, Hesse, Austina, & Oestrich, 2008; Osman, Barrios, Aukes, Osman, & Markway, 1993; Osman, Hoffman, Barrios, Kopper, Breitenstein, & Hahn, 2002).

- BSPS

The Brief Social Phobia Scale was found to be suitable to assess social phobia by Veale (2003). Davidson, Miner, De Veaugh-Geiss, Tupler, Colket and Potts (1997) concluded that the BSPS is a reliable, valid, and sensitive measure for the evaluation of social phobia. Nation-specific editions of the test have also been found to exhibit good psychometric properties (Osório, Crippa & Loureiro, 2010).

- SCL-90-R

The Symptoms Checklist-90-R has a long history of use for assessing psychopathology and co-morbidity screening (Sander and Jux, 2006), and also for alcohol-dependent patients (Choquette, 1994). SCL-90-R has been used to estimate relapse risk of alcoholics (Lucht, Jahn, Barnow, & Freyberger, 2002). Although a high intercorrelation between the factors of the inventory has been found (Zack, Toneatto,

& Streiner, 1998), the Global Severity Index score is considered a relevant predictor of relapse (Sander & Jux, 2006).

Validity and reliability of substance abuse measurement

- SMAST

The SMAST test has been validated by Pokorny, Miller and Kaplan (1972), Selzer, Vinokur and Van Roouen (1975), Zung (1979) and Zung (1984). A generally satisfying validity and reliability has been found in a variety of populations. However, Zung (1984) notes that previous validation studies may have been ”overly optimistic”.

- DAST-20

In a recent study, Cassidy and colleagues (2008) found the DAST-20 to perform well in the patient group investigated, with acceptable sensitivity and specificity. This is in line with the earlier findings of Gavin, Ross and Skinner (1989), who found a close agreement between DAST score and DSM-III diagnosis.

These measures of test-retest reliability here presented must be interpreted with caution. The time span between the ratings is long, and in that time span an intensive intervention has taken place.

External validity

The ability to generalise findings to other groups outside the experimental group is limited in this study. Substance-dependent people constitute a variety of human beings, with a large range in e.g. dependence history, previous and present fitness levels. Our subjects volunteered to take part, and this may represent a bias in the group, because people more interested in physical activity presumably would be more willing to participate. Such people might also initially be fitter, or less marked by chemical abuse. We can therefore not claim representativity for the subjects in our project,

making it difficult to generalise our findings to the wider substance-dependent population.

Reliability of physiological measurements

The trustworthiness of a measurement is reflected in the reliability of the measure. To ensure a high reliability, the metabolic analyser was frequently calibrated. The full calibration process consisted of a two-point gas calibration, volume calibration and calibration of barometric pressure and is described in more detail in Mamen, Resaland, Mo and Andersen (2008). The analyser has been compared to other analysers and found to behave consistently (Medbø et al., 2000).

The lactate analyser was calibrated with a "check strip" each test day, and with a "calibration strip" with each new package of strips. The performance of this analyser has also been extensively investigated and found to be reliable (Medbø et al., 2002). All the physiological post-test measures, MAP, utilisation, threshold cycling power and threshold running speed correlated strongly with their respective baseline measures ($p<0.006$, $r> 0.61$).

The speed of the treadmill was measured once a year by making a clearly visible line across the treadmill band and counting the times it passed during one minute at different speeds. The inclination was similarly checked for accuracy once a year. The Monark 824E ergometer cycle is self-calibrating and does not need any special calibration procedure. However, we did check from time to time that the friction threads attached to the weight basket did not cross, as this would have increased the friction.

5.5.3 Adherence to the programme and drop-outs

The adherence to the programme has been high, considering the large amount of training performed. Having a training partner was experienced as a valuable help to the participants, according to the evaluation at the end of the project. High adherence is a strength for the internal validity, but represents a threat to the external validity, as other studies have revealed less adherence to the training programme, and the use of training partners has not become the standard in training studies.

This study initially had a modest number of subjects (55 in total), and eight left before the baseline testing. We have no information about why they left, and know only their gender, age and recruiting source. Project drop-outs represent a bias hazard, as the ones who leave may have certain particular characteristics not always immediately apparent to the researcher or readily analysable with the data at hand.

Intention to treat

One way to compensate for this bias is to use **Intention to treat** analyses. Here all included participants are used in the evaluation, and for the drop-outs the post-treatment results equal the baseline values. In treatment research, persons who do not respond to the treatment may drop out, thereby creating an artificially beneficial effect of the treatment. Intention to treat analysis can therefore heighten the internal validity of the results, as the unknown bias is removed. On the other hand, assuming a zero development for the drop-outs is not always correct. In a training study, there should be some progress if the subject follows the instructions. Thus, in our case, a person who has experienced a positive gain in fitness and/or mental distress may not be motivated to stay in the project, and drop out.

5.5.4 Ethical considerations

The Helsinki declaration placed much emphasis on the protection of vulnerable groups. An individual may be considered vulnerable if he or she is dependent on treatment, and so might feel pressured to accept a suggestion from those who provide this treatment. This pressure was a concern in the present project. Information about voluntariness, and about alternative treatments were therefore strongly emphasized.

The distinction between research and treatment is not always clear, and those responsible for a project must make sure that the participants are fully informed of what is going to happen, even if the project is not considered a research project. This requirement of enhanced information was met in this project, as all participants received a comprehensive explanation of the project before signing the informed consent.

The training partners had to sign a declaration of confidentiality, so that their participants would not be exposed as substance abusers when taking part in training with other, healthy people.

Lastly, the lack of a control group also has an ethical aspect. The type of study design used limits the amount of new knowledge that could have been extracted from the project. It is therefore ethically questionable to expose the participants to long-term treatment without being fully able to harvest the evidence-based knowledge generated by the project. Although this is a major objection, it should also be said that the participants received much positive attention and learned to master new skills through their participation. Even with this simple set-up, it is possible to extract useful information about the project, to the benefit of the participants and others.

It should in addition also be noted that some control conditions, such as wait-lists and minimal interventions, may be suboptimal in terms of potential treatment gains. Thus, in some instance control conditions may represent unethical differential treatment between study participants.

5.5.5 Limitations

The economical limits render the use of a control group impossible, and this of course severely restricts the conclusions possible to make from the project, particularly in terms of cause and effect relationships.

The study uses research questions instead of hypothesis. The financial constraints made the use of a control group impossible, and this of course severely restricts the conclusions possible to make from the project, particularly in terms of cause and effect relationships. The initial number of subjects was modest, and the drop-out number substantial. The study design was simple, with no control group, which makes it impossible to claim causality. The screening of substance dependence was done by self-report, which makes the results vulnerable for manipulation, and no formal diagnoses were made. The sensitivity of these substance abuse tests is not very high, so changes in substance dependence can occur without the test being able to detect them. More suitable tests could have included the AUDIT and the DUDIT inventories, as they may be more sensitive, albeit also self-reported, inventories (Babor, Higgins-Biddle, Saunders, & Monteiro, 2001; Skinner, 1982). Consequently, the results should be interpreted with caution.

5.5.6 Strengths

Among the strengths of the study are the use of directly measured maximal aerobic power and measurements of lactate threshold. These are among the most accurate measures of physical fitness and those most sensitive to changes over time. The long training period and the use of training partners also represent strengths of the study. Enhancing aerobic power may take two to three months, but developing a new lifestyle where training plays an important part may take substantially longer. The training partners functioned as personal trainers, motivating and educating the participants to train regularly at suitable intensity. The choice of activities was large, and included both individual and group activities. All in all, this makes the project unique, and the experiences gained here may benefit future studies on training of substance-dependent people.

5.5.7 Implications for clinical practice and future research

This study supports the findings of previous studies, indicating that exercise interventions are beneficial for patients with substance-related problems and additional mental distress. These patients are normally sedentary, and exercise should therefore be more often incorporated in treatment programmes for such patients, to improve mental as well as physical health. The use of training partners was favourably rated by the participants and may be an important tool for programme adherence. The present study represents an improvement on previous studies by adopting precise measures of aerobic fitness. Future studies with control groups, preferably with a randomised, controlled design, are needed to obtain more valid information about the effect of exercise intervention in these patient groups. A follow up of this group would also be

very interesting, both with regards to degree of substance abuse, mental distress and training practice.

5.6 Conclusions

- Substance-dependent people can accomplish systematic training for a long time if assisted by a training partner. The development in physical fitness was to a large extent mimicked by the mental distress development: Those with larger fitness gains experienced greater improvements in mental distress.
- The MAP of chemically dependent persons, based on a maximal exertion test and directly measured oxygen uptake was 39.0 (10.0) $ml \bullet kg^{-1} \bullet min^{-1}$ for males and 30.8 (7.9) $ml \bullet kg^{-1} \bullet min^{-1}$ for females at project start. This is below the average for age- matched healthy people (Paper I).
- The improvement of MAP among the participants was about 5%. Submaximal lactate threshold performance rose about 11% (Paper II).
- The development of mental distress was related to the development of physical fitness, as those who improved their MAP also showed improvements in more mental distress issues than those who did not improve MAP (Paper III).
- Screening for substance abuse did not change during the project (Paper III)
- The training partners were positively evaluated by the project participants (Paper III)

6 References

Anderson, P., & Baumberg, B. (2006). *Alcohol in Europe. A public health perspective*. London: Institute of alcohol studies.

Armstrong, T. D., & Costello, E. J. (2002). Community Studies on Adolescent Substance Use, Abuse, or Dependence and Psychiatric Comorbidity. *Journal of Consulting & Clinical Psychology, 70*, 1224–1240.

Babor, T. F., Higgins-Biddle, J. C., Saunders, J. B., Monteiro, M. G. (2001). *The Alcohol Use Disorders Identification Test. Guidelines for Use in Primary Care. 2nd Ed.* Geneva: World Health Organization, Department of Mental Health and Substance Dependence.

Berman, A.H., Bergman, H., Palmstierna, T. , & Schlyter, F. (2005). Evaluation of the Drug Use Disorders Identification Test (DUDIT) in Criminal Justice and Detoxification Settings and in a Swedish Population Sample. *European Addiction Research,11*, 22–31. doi: 10.1159/000081413

Beck, A.T., & Steer, R.A. (1990). *Manual for the Beck Anxiety Inventory*. San Antonio, TX: Psychological Corporation.

Beck, A.T., & Steer, R.A. (1993). *Manual for the Beck Anxiety Inventory*. San Antonio: Psychological Corporation Harcourt Brace & Company.

Beck, A.T, Steer, R.A., & Brown, G.K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.

Beck, A., Wright, F.D., Newman, C.F., & Liese, B.S. (1993). *Cognitive therapy of substance abuse* (p. 4). New York: Guilford Press.

Biddle, S.J.H., & Mutrie, N. (2001). *Psychology of physical activity: determinants, well-being, and interventions*. London: Routledge.

Bijl, R.V, Ravelli, A., van Zessen, G. (1998). Prevalence of psychiatric disorder in the general population: results of The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Social Psychiatry and Psychiatric Epidemiology, 33*, 587–595.

Borg, G. (1962). *Physical performance and perceived exertion*. Thesis. Lund: Gleerup.

Borg G. (1970). Rating of perceived exertion as an indicator of somatic stress. *Scandinavian Journal of Rehabilitation Medicine*, 2, 92–98.

Borg G. (1982). Psychophysical bases of perceived exertion. *Medicine and Science in Sport and Exercise, 14*, 377–381.

Borg, G. (1998). *Borg's perceived exertion and pain scales*. Champaign, IL: Human Kinetics

Buckworth, J., & Dishman, R.K. (2002). *Exercise Psychology* (p. 14). Champaign: Human Kinetics.

Cassidy, C. M., Schmitz, N., & Malla, A. (2008). Validation of the Alcohol Use Disorders Identification Test and the Drug Abuse Screening Test in First Episode Psychosis. *The Canadian Journal of Psychiatry, 53*, 26–33.

Chalmers, R.J., Sulaiman, W.R., & Johnson, R.H. (1977). The metabolic response to exercise in chronic alcoholics. *Quarterly Journal of Experimental Physiology and Cognate Medical Sciences, 62*, 265–274.

Chen, M.J., Fan, X., & Moe, S.T. (2002). Criterion-related validity of the Borg ratings of perceived exertion scale in healthy individuals: a meta-analysis. *Journal of Sports Sciences*, *20,* 873–899.

Choquette, K.A. (1994). Assessing depression in alcoholics with the BDI, SCL-90R, and DIS criteria. *Journal of Substance Abuse, 6,* 295–304.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ, Lawrence Earlbaum Associates.

Collingwood, T. R., Reynolds, R., Kohl, H. W., Smith, W., & Sloan, S. (1991). Physical fitness effects on substance abuse risk factors and use patterns. *Journal of Drug Education, 21,* 73–84.

Collingwood, T.R., Sunderlin, J., Reynolds, R. & Kohl, H.W. (2000). Physical training as a substance abuse prevention intervention for youth. *Journal of Drug Education, 30*, 435–451.

Cronbach, L. J., &. Meehl, P. E. (1955). Construct Validity In Psychological Tests. *Psychological Bulletin*, *52*, 281–302.

Davidson, J. R. T., Miner, C. M., De Veaugh-Geiss, J., Tupler, L. A., Colket, J. T., & Potts, N. L. S. (1997). The Brief Social Phobia Scale: A Psychometric Evaluation. *Psychological Medicine. 27*, 161–166.

Davidson, J.R., Potts, N.L., Richichi, E.A., Ford, S.M., Krishnan, K.R., Smith, R.D., & Wilson, W. (1991). The Brief Social Phobia Scale. *The Journal of Clinical*

Psychiatry, 52, 48–51.

De Lorme, T. L. & Watkins, A. L. (1948). Technics of progressive resistance exercise. *Archives of Physical Medicine,* 264–273.

Derogatis, L.R. (1983). *Symptom Checklist-90 Revised: Administration, Scoring and Procedures Manual*. Minneapolis: National Computer Systems.

Donaghy, M.E., & Mutrie, N. (1999). Is exercise beneficial in the treatment and rehabilitation of the problem drinker? A critical review. *Physical Therapy Reviews, 4*, 153–66.

Donaghy, M.E., & Ussher, M.H. (2005). *Exercise interventions in drug and alcohol rehabilitation.* In: G.E.J Faulkner, & A.H Taylor (Eds.), *Exercise, health and mental health: emerging relationships* (pp 48–69). London: Routledge.

Enns, M.W., Cox, B.J., Parker, J.D.A., & Guertin, J.E. (1998). Confirmatory factor analysis of the Beck Anxiety and Depression Inventories in patients with major depression. *Journal of Affective Disorders, 47,* 195–200.

Ermalinski, R., Hanson, P.G., Lubin, B., Thornby, J.L. & Nahormek, P.A. (1997). Impact of a body-mind treatment component on alcoholic inpatients. *Journal of Psychosocial Nursing Mental Health Services, 35*, 39–45.

Esteve-Lanao, J., Foster, C., Seiler, S., & Lucia, A. (2007). Impact of training intensity distribution on performance in endurance athletes. *Journal of Strength and Conditioning Research, 21,* 943-94.

Faude, O., Kindermann, W., & Meyer, T. (2009). Lactate Threshold Concepts. How valid are they? *Sports Medicine, 39*, 469–490.

Fernandez-Sola, J., Estruch, R., Grau, J., Pare, J.C., Rubin, E., Urbano-Marquez, A. (1994). The Relation of Alcoholic Myopathy to Cardiomyopathy. *Annals of Internal Medicine, 120*, 529–536.

Fernandez-Solà, J., Preedy, V.R., Lang, C.H., Gonzalez-Reimers, E., Arno, M., Lin, J.C., Wiseman, H., Zhou, S., Emery, P.W., Nakahara, T., Hashimoto, K., Hirano, M., Santolaria-Fernández, F., González-Hernández, T., Fatjó, F., Sacanella, E., Estruch, R., Nicolás, J.M., Urbano-Márquez, A. (2007). Molecular and cellular events in alcohol-induced muscle disease. *Alcoholism, Clinical and Experimental Research, 31*, 1953–1962.

Fiskerstrand, Å. & Seiler, K. S. (2004). Training and performance characteristics

among Norwegian International Rowers 1970–2001. *Scandinavian Journal of Medicine & Science in Sports, 14*, 303–310.

Fitts, R.H. (2008). The cross-bridge cycle and skeletal muscle fatigue. *Journal of Applied Physiology, 104*, 551–558.

Foster, C., Fitzgerald, D.J., & Spatz, P. (1999). Stability of the blood lactate-heart rate relationship in competitive athletes. *Medicine & Science in Sports & Exercise, 31*, 578–582.

Frances, R.J., Miller, S.I., Mack, A.H. (2005). *Clinical Textbook of Addictive Disorders*. New York: The Guilford Press.

Frankel, A., Murphy, J. (1974). Physical fitness and personality in alcoholism. Canonical analysis of measures before and after treatment. *Quarterly Journal of Studies on Alcohol, 35,* 1272–1278.

Gavin, D.R., Ross, H.E., & Skinner, H.A. (1989). Diagnostic Validity of the Drug Abuse Sereening Test in the Assessment of DSM-III Drug Disorders. *British Journal of Addiction, 84*, 301–307.

Gary, V. & Guthrie, D. (1972). The effect of jogging on physical fitness and self-concept in hospitalized alcoholics. *Quarterly Journal of Studies on Alcohol, 33*, 1073–1078.

Gillis, M.M., Haaga, D.A.F., & Ford, G.T. (1995). Normative values for the Beck Anxiety Inventory, Fear Questionnaire, Penn State Worry Questionnaire, and Social Phobia and Anxiety Inventory. *Psychological Assessment, 7,* 450–455.

Goffman, E. (1959). *The presentation of self in everyday life*. Edinburgh: Edinburgh University.

Golding, L.A., Myers, C.R., Sinning, W.E. (1989). *Y's way to physical fitness the complete guide to fitness testing and instruction. 3rd ed.* (pp.114–120). Champaign, IL; Published for YMCA of the USA by Human Kinetics Publishers.

Goodwin, M.L., Harris, J.E., Hernández, A., & Gladden, L.B. (2007). Blood Lactate Measurements and Analysis during Exercise: A Guide for Clinicians. *Journal of Diabetes Science and Technology, 1*, 558–569.

Halson, S. L. & Jeukendrup, A.E. (2004). Does Overtraining Exist? An Analysis of Overreaching and Overtraining Research. *Sports Medicine, 34*,967–981.

Hansen, G., Blanchard, C., Rodgers, W., & Bell, G. (2003). Efficacy of Prescribing Endurance Training Intensity Using the Ventilatory Equivalents for Oxygen and Carbon Dioxide in Untrained Men and Women. *Research in Sports Medicine, 11*, 23–32.

Helgerud, J., Ingjer, F., & Strømme, S.B.(1990). Sex differences in performance-matched marathon runners. *European Journal of Applied Physiology, 61*, 433–439.

The Helsinki Declaration (http://www.cirp.org/library/ethics/helsinki/). Access date 23rd of September, 2010.

Hermansen, L., & Saltin, B. (1969). Oxygen uptake during maximal treadmill and bicycle exercise. *Journal of Applied Physiology, 26*, 31–37.

Hesse, M. (2006). The Beck Depression Inventory in patients undergoing opiate agonist maintenance treatment. *British Journal of Clinical Psychology, 45*, 417–425.

Hewitt, P.L., & Norton, G.R. (1993). The Beck Anxiety Inventory: A psychometric analysis. *Psychological Assessment, 5,* 408–412.

Jakovlev, N. N. (1977). *Sportbiochemie, Sportmedizinisches Schriftenreihe, Band 14*. Leipzig: Johann Ambrosius Barth.

Jones, S.R.G. (1992). Was there a Hawthorne effect? *The American Journal of Sociology, 3*, 451–468.

Jørgensen, K. (1985). Permissible loads based on energy expenditure measurements. *Ergonimics, 28*, 365–369.

Kershaw, C.D. & Guidot, D.M. (2008). Alcoholic lung disease. *Alcohol Research & Health, 31*, 66–75.

Kuipers, H. (1983). *Variability of physiological responses to exercise*. Thesis. Haarlem: Uitgeverij de Vriesborch.

Lucía, A., Morán, M., Zihong, H., Z., & Ruiz, J.R. (2010). Elite Athletes: Are the Genes the Champions? *International Journal of Sports Physiology and Performance, 5*, 98–102.

Lucht, M., Jahn, U., Barnow, S., & Freyerberger, H.J. (2002). The use of a symptom checklist (SCL-90-R) as an easy method to estimate the relaps risk after alcoholism detoxification. *European Addiction Research, 8*, 190–194.

DOI:10.1159/000066131

Lundin, L., Hällgren, R., Landelius, J., Roxin, L.-E., Venger, P. (1986). Myocardial and skeletal muscle function in habitual alcoholics and its relation to serum myoglobin. *American Journal of Cardiology, 58*, 795–799.

Lykke, J., Hesse, M., Austina, S.F., & Oestrich, I. (2008). Validity of the BPRS, the BDI and the BAI in dual diagnosis patients. *Addictive Behaviors, 33,* 292–300.

Mamen, A., Resaland, G.K., Mo, D.A., Andersen, L.B. (2008). Comparison of peak oxygen uptake in boys exercising on treadmill and cycle ergometers. *Gazzetta Medica Italiana Archivio per le Scienze Mediche, 167*, 15–21.

Mamen, A., Skrede, A., Munkvold, H., & Martinsen, E.W. (2009). Rating of perceived exertion in substance dependent adults with attention deficit/hyperactivity disorder: An issue of self presentation? *Gazzetta Medica Italiana Archivio per le Scienze Mediche, 168*, 121–128.

Martin, J.E., & Dubbert, P.M. (1982). Exercise Applications and Promotion in Behavioral Medicine: Current Status and Future Directions. *Journal of Consulting and Clinical Psychology, 50*, 1004–1017.

Martinsen, E.W. (2008). Physical activity in the prevention and treatment of anxiety and depression. *Nordic Journal of Psychiatry, 62* (Suppl 47), 25–29.

Martinsen, E.W., Raglin, J.S. (2007). Anxiety/Depression: Lifestyle medicine approaches. *American Journal of Lifestyle Medicine, 1*, 159–166.

Massie, J.F. & Shephard, R.J. (1971). Physiological and psychological effects of training: a comparison of individual and gymnasium programs, with a characterization of the exercise "drop-out". *Medicine and Science in Sports, 3*, 110–117.

McArdle W.D, Katch F.I, & Katch V.L. (2010). *Exercise physiology energy, nutrition, and human performance. 7th ed.* Baltimore; Walter Kluwer, Williams & Wilkins.

McLean, S. R., Norris, S. R., & Smith, D. J. (2004). Comparison of the Lactate Pro and the YSI 1500 Sport Blood Lactate Analyzers. *International Journal of Applied Sports Sciences, 16,* 22–31.

Medbø, J. I., Mamen, A., Holt Olsen, O., & Evertsen, F. (2000). Examination of four different instruments for measuring blood lactate concentration. *Scandinavian*

Journal of Clinical and Laboratory Investigation, 60, 367–380.

Medbø, J.I., Mamen, A., Welde, B., von Heimburg, E., & Stokke, R. (2002). Examination of the Metamax I and II oxygen analysers during exercise studies in the laboratory. *Scandinavian Journal of Clinical & Laboratory Investigation, 62,* 585–598.

Melberg, H. O. (2005*). Costs and benefits of treating drug users: Essays on selection bias, contagious drug use, the gateway hypothesis and the concept of social cost* (Doctoral dissertation, University of Oslo, Faculty of Social Sciences, Norway).

Midgley, A. W., McNaughton, L. R. & Wilkinson, M. (2006). Is there an Optimal Training Intensity for Enhancing the Maximal Oxygen Uptake of Distance Runners? *Sports Medicine, 36*, 117–132.

Morris, J.N., Heady, J.A., Raffle, P.A., Roberts, C.G., Parks, J.W. (1953). Coronary heart-disease and physical activity of work. Lancet, 265, 1053–1057.

Mukherjee, S., & Chia, H.M.Y. (2006). Evaluation of the Lactate Pro portable blood-lactate analyser involving multiple-tester approach. *Asian Journal of Exercise & Sports Science, 3*, 55–60.

Murphy, J.B. (1970). An approach to the treatment of alcoholism through corrective therapy. *American Corrective Therapy Journal, 24*, 88–92.

Murphy, T., Pagano, R. R., & Marlatt, G. A. (1986). Lifestyle modification with heavy alcohol drinkers: effects of aerobic exercise and meditation. *Addictive Behaviors, 11,* 175–186.

Noble, B.J. & Robertson, R.J. (1996). *Perceived Exertion.* Champaign, IL: Human Kinetics.

Nordström, A. & Bodlund, O. (2008). Every third patient in primary care suffers from depression, anxiety or alcohol problems. *Nordic Journal of Psychiatry, 62*, 250–255.

The Norwegian Institute of Public Health, http:// www.fhi.no/eway/default.aspx?pid=233&trg=MainArea_5661&MainArea_5661=5631:0:15,2667:1:0:0:::0:0.

Osman, A., Barrios, F.X., Aukes, D., Osman, J.R., & Markway, K. (1993). The Beck Anxiety Inventory: psychometric properties in a community population.

Journal of Psychopathology and Behavioral Assessment, 15, 287–297.

Osman, A., Hoffman, J., Barrios, F.X., Kopper, B.A., Breitenstein, J.L., & Hahn, S.K. (2002). Factor structure, reliability, and validity of the Beck Anxiety Inventory in adolescent psychiatric inpatients. *Journal of Clinical Psychology, 58,* 443–456.

Osório, F.L., Crippa, J.A.S., & Loureiro, S.R.(2010). Study of the psychometric qualities of the Brief Social Phobia Scale (BSPS) in Brazilian university students. *European Psychiatry, 25*, 178–188. doi: 10.1016/j.eurpsy. 2009.08.002

Palmer, J., Vacc, N., & Epstein, J. (1988). Adult inpatient alcoholics: physical exercise as a treatment intervention. *Journal of Studies on Alcohol, 49,* 418–421.

Paffenbarger, R.S. Jr, Wing, A.L., & Hyde, R.T. (1978). Physical activity as an index of heart attack risk in college alumni. *American Journal of Epidemiology, 108*, 161–175.

Pokorny, A.D., Miller, B.A., & Kaplan, H.B. (1972). The brief MAST: a shortened version of the Michigan Alcoholism Screening Test. *The American Journal of Psychiatry, 129*, 342–345.

Purvis, J.W. & Cureton, K.J. (1981). Ratings of perceived exertion at the anaerobic threshold. *Ergonomics, 24*, 295–300.

Read, J. & Brown, R.A. (2003). The role of physical exercise in alcoholism treatment and recovery. *Professional Psychology: Research and Practice, 34*, 49–56.

Read, J.P., Brown, R.A., Marcus, B.H., Kahler, C.W., Ramsay, S.E., Dubreuil, R.N., Jakicic, J.M., & Francione, C. (2001). Exercise attitudes and behaviours among persons in treatment for alcohol use disorders. *Journal of Substance Abuse Treatment, 21*, 199–206.

Regier, D.A., Farmer, M.E., Rae, D.S., Locke, B., Z., Keith, S.J., Judd, L.L., Goodwin, F.K. (1990). Comorbidity of mental disorders with alcohol and other drug abuse. *The Journal of the American Medical Association, 264*, 2511–2518.

Raymer, G.H., Marsh, G.D., Kowalchuk, J.M, & Thompson, R.T. (2004). *Journal of Applied Physiology, 96*, 2050–2056.

Richter, P., Werner, J., Heerlein, A., Kraus, A., & Sauer, H. (1998). On the validity of the Beck Depression Inventory. A review. *Psychopathology 31*, 160–168.

Rowland, T. (2009). Endurance Athletes' Stroke Volume Response to Progressive Exercise. A Critical Review. *Sports Medicine, 39*, 687–695.

Rubin, E. (1979). Alcoholic myopathy in heart and skeletal muscle. *The New England Journal of Medicine, 301,* 28–33.

Russo, C. A. & Elixhauser, A. (2006). *Hospitalizations for Alcohol Abuse Disorders, 2003.* HCUP Statistical Brief #4. Agency for Healthcare Research and Quality, Rockville, MD. http://www.hcup-us.ahrq.gov/reports/statbriefs /sb4.pdf.

Sander,W., Jux, M. (2006). Psychological distress in alcohol-dependent patients evaluating inpatient treatment with the Symptom Checklist (SCL-90-R). *European Addiction Research, 12*, 61–66.

Sanderson, B.K., Phillips, M.M., Gerald, L., DiLillo, V., & Bittner, V. (2003). Factors associated with the failure of patients to complete cardiac rehabilitation for medical and nonmedical reasons. *Journal of Cardiopulmonary Rehabilitation, 23*, 281–289.

Segerström, A. B., Glans, F., Eriksson, K. F., Groop, L., Thorsson, O., & Wollmer, P. (2008). Assessment of exercise capacity in women with type 2 diabetes. *Clinical Physiology and Functional Imaging, 28,* 294–298.

Seiler, K.S. & Kjerland, G.O. (2006). Quantifying training intensity distribution in elite endurance athletes: Is there evidence for an 'optimal' distribution? *Scandinavian Journal of Medicine and Science in Sport, 16,* 49–56.

Sell, H. & Christensen, N. J. (1989). Fysisk trænings effekt på fysiske, psykiske og sociale forhold hos stof- og/eller alkoholmisbrugere [The effect of physical training on physical, mental and social conditions in drug and/or alcohol addicts. In Danish.] *Ugeskrift for læger, 151,* 2064–2067.

Statistics Norway (www.ssb.no)

Selzer, M.L., Vinokur, A., & van Rooijen, L. (1975). A self-administered short version of the Michigan Alcoholism Screening Test (SMAST). *Journal of Studies on Alcohol, 36*, 117–126.

Sesso, H. D., Paffenbarger Jr, R.S., &, Lee, I-M. (2000). Physical Activity and Coronary Heart Disease in Men: The Harvard Alumni Health Study. *Circulation ,102*, 975-980.

Sexton, H, Mære, A., & Dahl, N.H. (1989).Exercise intensity and reduction in neurotic

symptoms. A controlled follow-up study. *Acta Psychiatrica Scandinavica, 80*, 231–235.

Sinyor, D., Brown, T., Rostant, L., & Seraganian, P. (1982). The role of a physical-fitness program in the treatment of alcoholism. *Journal of Studies on Alcohol, 43,* 380–386.

Skinner, H.A. (1982). The drug abuse screening test. *Addictive Behaviors, 7*, 363–371.

Skaset, H.B. (1976). *Idrettsterminologi* [Sports terminology]. Rud, Norges Idrettsforbund.

Skrede, A., Munkvold, H., Watne, Ø., & Martinsen, E.W. (2006). Treningskontaktar ved rusproblem og psykiske lidingar [Use of local exercise contacts in the treatment of substance dependence and mental disorders, in New Norwegian with English abstract]. *Tidsskrift for Norsk Lægeforening, 126*, 1925–1927.

Stathopoulou, G., Powers, M.B., Berry, A.C., Smits, J.A.J. & Otto, M.W. (2006). Exercise interventions for mental health: A quantitative and qualitative review. *Clinical Psychology: Science in Practice, 13*, 179–193.

Timmons, J.A., Knudsen, S., Rankinen, T., Koch, L.G., Sarzynski, M., Jensen, T., Keller, P., Scheele, C., Vollaard, N.B.J., Nielsen, S., Åkerström, T., MacDougald, O. A., Jansson, E., Greenhaff, P.L., Tarnopolsky, M.A., van Loon, L.J.C., Pedersen, B.K., Sundberg, C.J., Wahlestedt, C., Britton, S. L., & Bouchard, C. (2010). Using molecular classification to predict gains in maximal aerobic capacity following endurance exercise training in humans. *Journal of Applied Physiology*, *108*, 1487–96.

Tjønna, A.E., Lund Nilsen, T.I., Slørdahl, S.A., Vatten, L., & Wisløff, U. (2010). The association of metabolic clustering and physical activity with cardiovascular mortality: the HUNT study in Norway. *Journal of Epidemiology and Community Health*, *64*, 690–695.

Tkachuk, G.A. & Martin, G.L. (1999). Exercise Therapy for Patients with Psychiatric Disorders: Research and Clinical Implications. *Professional Psychology - Research & Practice. 3*, 275–282.

Urbano-Marquez, A., Estruch, R., Navarro-Lopez, F., Grau, J.M., Mont, L. & Rubin, E. (1989). The effect of alcoholism on skeletal and cardiac muscle. *The New England Journal of Medicine, 320*, 409–415.

Veale, D. (2003). Treatment of social phobia. *Advances in Psychiatric Treatment, 9,* 258–264.

Vollaard, B.J. Constantin-Teodosiu, D., Fredriksson, K., Rooyacker, O. Jansson, E., Greenhaff, P. L., Timmons, J.A., & Sundberg, C. J. (2009). Systematic analysis of adaptations in aerobic capacity and submaximal energy metabolism provides a unique insight into determinants of human aerobic performance. *Journal of Applied Physiology, 106,* 1479–1486.

Weltman, A. (1995).*The blood lactate response to exercise.* Champaign, IL: Human Kinetics.

Wenger, H.A. & Bell, G.J. (1986). The interactions of intensity, frequency and duration of exercise training in altering cardio-respiratory fitness. *Sports Medicine, 3*, 346–356.

Westermeyer, J. (2005). Historical and social context of psychoactive substance disorders. In: R.J. Frances, & S.I. Miller (Eds.), *Clinical textbook of addictive disorders* (pp. 16–25). New York: Guilford.

Wisløff, U., Ellingsen, Ø., & Kemi, O.J. (2009). High-Intensity Interval Training to Maximize Cardiac Benefits of Exercise Training? *Exercise and Sport Sciences Reviews, 37*,139–146.

Zack, M., Toneatto, T., & Streiner, D.L. (1998). The SCL-90 factor structure in comorbid substance abusers. *Journal of Substance Abuse, 10*, 85–101.

Zung, B,J. (1979). Psychometric properties of the Michigan Alcoholism Screening Test and briefer versions. *Journal of Studies on Alcohol, 40,* 845–859.

Zung, B. J. (1984). Reliability and validity of the short MAST among psychiatric inpatients. *Journal of Clinical Psychology, 40*, 347–350.

Åstrand, P.-O., Rodahl, K., Dahl, H.A., & Strømme, S.B. (2003). *Textbook of work physiology, 4th ed.* Champaign, IL: Human Kinetics.