Association between depression and cardiorespiratory fitness in general population and patients with heart disease

Theodoros Papasavvas

A thesis submitted for the degree of PhD

School of Sport, Rehabilitation and Exercise Sciences
University of Essex
March 2019
To father and mother
Acknowledgments

This project would have not been possible without the support and guidance of my supervisors, Prof. Dominic Micklewright and Dr. Gavin Sandercock, whom I deeply thank. I would also like to acknowledge the contribution of Dr. Mohommad Alhashemi, Prof. Robert Bonow, Prof. Hassen Alamin, and Dr. Hanny Ghabrash in the papers that they co-authored with me. Last, I am grateful to Nermin, Rasha, and Reem for their help in collecting data for this project.
Abstract

Depression symptom severity and cardiorespiratory fitness (CRF) are both predictors of mortality and disability in healthy individuals and patients with heart disease. However, the relationship between the two is unclear. We conducted two systematic reviews and meta-analyses in otherwise healthy individuals and patients with heart disease, respectively, in order to assess the relationship between depression symptom severity and CRF. The first study (Papasavvas et al., 2016b) included 16 studies (4039 participants) and revealed a negative correlation between depression symptom severity and CRF [correlation coefficient (CC -0.16, 95 % CI -0.21 to -0.10)] that appeared stronger in male (CC -0.22, 95 % CI -0.26 to -0.18) than in female individuals (CC -0.12, 95 % CI -0.19 to -0.05; p = 0.01). The second study (Papasavvas et al., 2017) included 59 studies (25733 participants) and also revealed a negative correlation between depression symptom severity and CRF (CC − 0.15; 95% CI, − 0.17 to − 0.12) that was independent to sex. Within-study level moderator analysis was not possible because raw data were not available for every study. I also assessed the effects of potential moderators of the correlation between depression symptom severity and CRF in 1489 patients with heart disease using linear and logistic regression analysis. Sex (p = 0.007) and BMI (p < 0.001) moderated the correlation: Lower BMI enhanced the correlation, while higher BMI attenuated the correlation in male patients and rendered it statistically not significant in female patients (Papasavvas 2018 - unpublished). We also translated the Cardiac Depression Scale to Arabic and validated it in a representative sample of 260 Arab patients with heart disease (Papasavvas et al., 2016a), in order to assess depression symptom severity in the following study. The subsequent quasi-experimental study aimed to assess whether depression reduces CRF in patients with heart disease but this was eventually not feasible. The above findings have clinical and prognostic implications and should stimulate further research on the effects of improving depression symptom severity on CRF and vice versa. Potential causative relationships between depression symptom severity and CRF should also be investigated.
Note on citation styles:

Each published paper follows the required citation style of the respective publisher. The remainder of the thesis follows the same citation style.
Table of contents

Acknowledgments 3
Abstract 4
**Chapter 1: Introduction** 8
A. Depressive symptoms and cardiorespiratory fitness in healthy individuals 9
   A.1 Epidemiology of depression 9
   A.2 Determinants of depression 10
   A.3 Burden of depression 12
   A.4 Exercise and cardiorespiratory fitness 12
   A.5 Determinants of cardiorespiratory fitness 14
   A.6 Relationship between depression and cardiorespiratory fitness 19
B. Depressive symptoms and cardiorespiratory fitness in patients with heart disease 21
   B.1 Epidemiology of depression in patients with heart disease 21
   B.2 The link between depression and heart disease 22
   B.3 Impact of depression on patients with heart disease 27
   B.4 Exercise and cardiorespiratory fitness in patients with heart disease 28
   B.5 Relationship between depression and cardiorespiratory fitness in patients with heart disease 29
C. Moderators of the relationship between depression and CRF in patients with heart disease 30
D. Translation and validation of the Cardiac Depression Scale to Arabic 30
E. Is the correlation between depression and cardiorespiratory fitness causative? 32
References 33

**Chapter 2: Is there an association between depression symptom severity and CRF in otherwise healthy individuals?** 52
Abstract 53
1 Introduction 54
2 Methods 56
3 Results 61
4 Discussion 70
5 Conclusions 75
References 76

**Chapter 3: Is there an association between depression symptom severity and CRF in patients with heart disease?** 81
Abstract 82
Introduction 83
Methods 84
Results 88
Discussion 101
Conclusions 104
References 105

**Chapter 4: Is the relationship between depression symptom severity and CRF in patients with heart disease moderated?** 110
Abstract 111
Introduction 112
Methods 114
Results 115
Chapter 1

Introduction
Introduction

A. Depressive symptoms and cardiorespiratory fitness in healthy individuals

A.1 Epidemiology of depression

Depression is one of the most common mental health disorders, characterized by persistent (lasting for at least two weeks) depressed mood, loss of interest or pleasure, significant weight loss or weight gain, feelings of guilt and worthlessness, fatigue, poor concentration, and recurrent thoughts of death, although these symptoms are not necessarily present altogether (American Psychiatric Association 2013). Depression is the leading cause of disability worldwide affecting 322 million people in the world (4.4% of global population), with females being more affected by depression than males (5.1% versus 3.6% of global population) (World Health Organization 2017). A recent meta-analysis using data from 30 countries reported one-year prevalence of depression of 7.2% (Lim et al., 2018). One-year prevalence of depression ranges between 4% - 4.9% in Africa (Tomlinson et al., 2009, Peltzer and Phaswana-Mafuya 2013), 4.8% - 8% in America (Patten et al., 2006, Slone et al., 2006, Silva et al., 2014), 1.7% - 6.7% in Asia (Chiu 2004, Lee et al., 2009), is 6.2% in Australia (Australian Institute of Health and Welfare 2018), and 6.7% in Europe (World Health Organization 2013).

Depression is more prevalent among patients. A recent meta-analysis reported that depression or depressive symptoms were present in 27% of outpatients (Wang et al., 2017). Depression is present in a wide array of patients and its prevalence ranges from 5% - 60%, depending on the underlying disease (see Table 1.1).
<table>
<thead>
<tr>
<th>Underlying disease</th>
<th>Prevalence of depression (%)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parkinson’s disease</td>
<td>35</td>
<td>(Aarsland et al., 2011)</td>
</tr>
<tr>
<td>Stroke</td>
<td>33</td>
<td>(Poynter et al., 2009)</td>
</tr>
<tr>
<td>Acute orthopedic trauma</td>
<td>32.6</td>
<td>(Muscatelli et al., 2017)</td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td>30.5</td>
<td>(Boeschoten et al., 2017)</td>
</tr>
<tr>
<td>Chronic obstructive pulmonary disease</td>
<td>27.1</td>
<td>(Matte et al., 2016)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>26.8</td>
<td>(Li et al., 2015)</td>
</tr>
<tr>
<td>Type 2 diabetes</td>
<td>25 – 38</td>
<td>(Semenkovich et al., 2015, Hussain et al., 2018)</td>
</tr>
<tr>
<td>Systemic erythematous lupus</td>
<td>24</td>
<td>(Zhang et al., 2017)</td>
</tr>
<tr>
<td>Chronic kidney disease</td>
<td>22.8</td>
<td>(Palmer et al., 2013)</td>
</tr>
<tr>
<td>Age-related macular degeneration</td>
<td>15.7 – 44</td>
<td>(Dawson et al., 2014)</td>
</tr>
<tr>
<td>Rheumatoid arthritis</td>
<td>14.8 – 38.8</td>
<td>(Matcham et al., 2013)</td>
</tr>
<tr>
<td>Alzheimer’s disease</td>
<td>12.7 – 42</td>
<td>(Chi et al., 2015)</td>
</tr>
<tr>
<td>Cancer</td>
<td>5 - 60</td>
<td>(Walker et al., 2013, Yang et al., 2013, Watts et al., 2014, Watts et al., 2015, Caruso et al., 2017)</td>
</tr>
</tbody>
</table>

Prevalence of depression in patients

**A.2 Determinants of depression**

Several models and theories exist aiming to describe potential causes of depression and explain their interactions. Kandel’s understanding of the interactions between mind and brain provides a template to study the nature of these causes (Kandel 1998); according to his model,

- all mental processes derive from the brain
- genes and their protein products determine neuronal connections and functioning
- life experiences influence gene expression and psychosocial factors feed back to the brain
- altered gene expression that produces changes in neuronal connections contributes to maintaining abnormalities of behavior

As it can be seen in the model, depression is caused by genetic and environmental factors, and their interaction. There is consensus that depression is twice more common in female than male individuals (Sloan...
and Kornstein 2003, Noble 2005). Twin studies have also suggested that about 40% of the susceptibility to depression is attributed to genes (Sullivan et al., 2000, Rice et al., 2002). Several environmental risk factors for the onset, relapse, and recurrence of depression have been proposed, including childhood abuse (Norman et al., 2012, Infurna et al., 2016, Targum and Nemeroff 2019), poverty (Beardslee et al., 2012, Yatham et al., 2018), negative family emotions and parental divorce (Di Manno et al., 2015, Aktar and Bogels 2017), and stressful life events (Hammen 2005, Yang et al., 2015). Depression has also been associated with an array of disorders (see Table 1.2).

**Table 1.2**

<table>
<thead>
<tr>
<th>Disorder</th>
<th>Indicative references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute orthopedic trauma</td>
<td>(Muscatelli et al., 2017)</td>
</tr>
<tr>
<td>Age-related macular degeneration</td>
<td>(Dawson et al., 2014)</td>
</tr>
<tr>
<td>Alzheimer’s disease</td>
<td>(Chi et al., 2015)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>(Pollack 2005)</td>
</tr>
<tr>
<td>Cancer</td>
<td>(Walker et al., 2013, Yang et al., 2013, Watts et al., 2014, Watts et al., 2015, Caruso et al., 2017)</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>(Celano and Huffman 2011, Dhar and Barton 2016)</td>
</tr>
<tr>
<td>Chronic kidney disease</td>
<td>(Palmer et al., 2013)</td>
</tr>
<tr>
<td>Chronic obstructive pulmonary disease</td>
<td>(Matte et al., 2016)</td>
</tr>
<tr>
<td>Chronic pain</td>
<td>(Humo et al., 2019)</td>
</tr>
<tr>
<td>Epilepsy</td>
<td>(Mula 2017)</td>
</tr>
<tr>
<td>HIV/AIDS</td>
<td>(Nanni et al., 2015)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>(Li et al., 2015)</td>
</tr>
<tr>
<td>Lyme Borreliosis</td>
<td>(Bransfield 2018)</td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td>(Boeschoten et al., 2017)</td>
</tr>
<tr>
<td>Obesity</td>
<td>(Quek et al., 2017, Ouakinin et al., 2018)</td>
</tr>
<tr>
<td>Parkinson’s disease</td>
<td>(Aarsland et al., 2011)</td>
</tr>
<tr>
<td>Porphyria</td>
<td>(Millward et al., 2005)</td>
</tr>
<tr>
<td>Rheumatoid arthritis</td>
<td>(Matcham et al., 2013)</td>
</tr>
<tr>
<td>Sleep disorders</td>
<td>(Riemann et al., 2019)</td>
</tr>
<tr>
<td>Stroke</td>
<td>(Poynter et al., 2009)</td>
</tr>
<tr>
<td>Syphilis</td>
<td>(Crozatti et al., 2015)</td>
</tr>
<tr>
<td>Systemic erythematous lupus</td>
<td>(Zhang et al., 2017)</td>
</tr>
<tr>
<td>Thyroid dysfunctions</td>
<td>(Williams et al., 2009)</td>
</tr>
<tr>
<td>Traumatic brain injury</td>
<td>(Rapoport 2012)</td>
</tr>
<tr>
<td>Type 2 diabetes</td>
<td>(Semenkovich et al., 2015, Hussain et al., 2018)</td>
</tr>
</tbody>
</table>

Disorders associated with depression
A.3 Burden of depression

Patients with depression have an increased risk of overall mortality of 52% - 107% (Cuijpers et al., 2014, Laursen et al., 2016, Brandao et al., 2018) and poorer global and generic health-related quality of life than non-depressed individuals (Sivertsen et al., 2015). Adding to this, non-adherence to treatment is very common (Pampallona et al., 2002) with only 25% - 50% of patients adhering to their antidepressant regimen for the length of time recommended by depression guidelines (Trivedi et al., 2007). The burden of depression, though, does not fall only on the patients but also on their families and care-givers, employers and insurance payers (Greenberg and Birnbaum 2005). There is economic burden, predominantly due to increased healthcare utilization and spending, absenteeism, and reduced productivity at work (Wang et al., 2003, Donohue and Pincus 2007) and there is a positive correlation between levels of treatment-resistant depression and direct and indirect costs (Johnston et al., 2019). The annual cost of this burden is estimated to be 332 - 370 dollars per capita in the USA (Mrazek et al., 2014) and 225 euros per capita in Europe (Olesen et al., 2012). Reasons for the heavier economic burden of depression in the USA compared to Europe may be differences in the above factors that constitute it between the two economies. Additionally, depression point prevalence seems to be higher in the USA than in Europe (13.4% vs 11.9%) (Lim et al., 2018), which could contribute to the higher economic burden in the former economy.

A.4 Exercise and cardiorespiratory fitness

Exercise as a tool to prevent and treat depression has been extensively studied in the last years. There is evidence that physical activity and exercise may help prevent depression (He et al., 2012, Lee et al., 2018, Pascoe and Parker 2018) even when individuals engage in minimal levels of one hour of physical activity each week (Harvey et al., 2018). Regarding treatment of depression, several meta-analyses have shown a favorable effect of exercise on reducing depression symptom severity across the age spectrum, with effect sizes ranging between 0.34 – 1.11 (Bridle et al., 2012, Cooney et al., 2013, Wegner et al., 2014, Kvam et al., 2016, Schuch
et al., 2016). Exercise has also been reported to improve physical and psychological quality of life in patients with depression (Schuch et al., 2016).

Engaging in regular exercise increases cardiorespiratory fitness (CRF), a well-established strong and independent predictor of all-cause and disease-specific mortality and disability (Castillo-Garzón et al., 2006, Harber et al., 2017), including cardiovascular disease, hypertension, dyslipidemia, diabetes, and atrial fibrillation (Laukkanen et al., 2002, Sui et al., 2017, Al-Mallah et al., 2018). There is general consensus that adults should engage in 150 minutes of moderate-intensity or 75 minutes of vigorous-intensity aerobic physical activity each week, and in muscle-strengthening activities involving major muscle groups two or more days each week (American Heart Association 2018, Australian Department of Health 2018, British National Health System 2018, World Health Organization 2018). However, 31% of the global adult population did not meet these recommendations in 2008 (World Health Organization 2014). Physical inactivity causes 3.2 million deaths each year (World Health Organization 2014) and costs 0.3% to 4.6% of national health expenditure (Ding et al., 2017).

Being physically inactive leads to low CRF, which is associated with approximately 70% higher risk for all-cause mortality compared to high CRF (Kokkinos et al., 2008, Kodama et al., 2009, Imboden et al., 2018) regardless of body mass index (BMI) (Barry et al., 2014). On the contrary, each 10% increase in CRF is associated with 16% decreased risk of all-cause mortality (Hussain et al., 2018). Moreover, high levels of CRF seem to attenuate the health risks of obesity (LaMonte and Blair 2006), as overweight fit individuals seem to have lower risk of mortality than normal-weight unfit individuals (Fogelholm 2010, Yerrakalva et al., 2015). The significance of CRF as a predictor of mortality and disability is also reflected by the recent American Heart Association’s advisory to treat CRF as a clinical vital sign (Ross et al., 2016).
A.5 Determinants of cardiorespiratory fitness

Similarly to the combination of genetic and environmental factors that predict depression, CRF is also determined by the interplay of nature and nurture. Genetic, environmental, and behavioral factors influence CRF (see Figure 1.1), each of which will be separately but briefly analyzed below.

![Determinants of cardiorespiratory fitness](image)

A.5.1 Genes

The contribution of genes to the variance of CRF is being studied since the 70s’; however, there is still no consensus on its magnitude. Studies on monozygotic and dizygotic twins have reported VO\(_{2}\)\text{MAX} to be genetically determined by as much as 93% in males (Klissouras 1971), 87% in males and 69% in females (Maes et al., 1996), 66% in males (Fagard et al., 1991), 27% adjusted for age and sex in the large HERITAGE Family Study (Bouchard et al., 1999), and even non-significant genetic contributions (Howald 1976). The large differences in these results can be attributed to methodological differences between studies, such as different
sample sizes and different or no adjustments for factors that influence CRF, and methodological considerations that might have significant impact on the result. In the study of Howald, for example (Howald 1976), the exclusion from the data analysis of two pairs of monozygotic twins whose co-twins engaged in intensive endurance training (and therefore were exposed to contrasting environments) resulted in 68% genetic determination of CRF. Although there is no consensus on the magnitude of genetic contribution to CRF, researchers agree that the contribution is significant (Georgiades et al., 2017, Ghosh et al., 2019).

A.5.2 Age

The effect of age on CRF has been studied since the early 50s’, where an inverse association between age and CRF was found in 7000 participants (Cullumbine et al., 1950). This finding has been confirmed multiple times and in large cohorts of 3250 (Wang et al., 2010), 57085 (Blaha et al., 2016), and 1169 participants (Cameron et al., 2018).

A.5.3 Gender

There is consensus that men tend to have higher CRF than women. This finding has been confirmed in studies including large samples (Kaminsky et al., 2015, Al-Mallah et al., 2016, Rossi Neto et al., 2019) and is consistent across the age spectrum (Hakola et al., 2011). The difference in CRF can be explained by physiological differences that affect CRF between genders, such as heart size (Hutchinson et al., 1991, Wingate 1997), lung volume (Bellemare et al., 2003), respiratory function (LoMauro and Aliverti 2018), and body mass composition (Bredella 2017), as well as by the fact that women may tend to be less physically active than men (Rockwood et al., 2004, Troiano et al., 2008). CRF also seems to increase more in men than in women in response to the same dose of exercise training (Diaz-Canestro and Montero 2019).
A.5.4 Physical activity

Self-reported physical activity levels are moderately correlated to CRF (Minder et al., 2014) and individuals who report a high level of physical activity in leisure time have higher CRF than individuals who are sedentary (Eriksen et al., 2016). These findings are confirmed in studies that measure physical activity objectively using accelerometers. Van der Velde et al. reported moderate correlations between objectively measured physical activity levels and CRF in a sample of 2024 adults (Van der Velde et al., 2017), while male and female individuals who met the physical activity recommendation of 150 min/week of daily moderate intensity physical activity had 13% and 9% higher VO$_{2\text{MAX}}$, respectively, than individuals who did not meet this recommendation (Dyrstad et al., 2016). Similarly, each additional hour of physical activity has been associated with a 0.88 metabolic equivalents of task (MET) higher fitness for men and a 1.37 METs higher fitness for women (Kulinski et al., 2014). A recent meta-analysis reported that workplace-based physical activity interventions can increase VO$_{2\text{MAX}}$ by 0.77 METs (Burn et al., 2019). It should be noted that physical activity can independently predict CRF even if it is conducted sporadically, although it should be of at least moderate intensity (McGuire and Ross 2011). It seems that physical activity volume is more important in predicting CRF than pattern or variability of intensities throughout the day (Knaeps et al., 2017).

A.5.5 Race

Studies including large samples, in which CRF was either estimated (Farrell et al., 1987, Ceaser et al., 2013, Pandey et al., 2016) or objectively measured using a symptom-limited exercise test (Sidney et al., 1992, Lavie et al., 2004, Al-Mallah et al., 2016), have reported that black individuals achieve a lower CRF compared to white individuals. Black men have also been reported to have more than twice the risk of being unfit compared to white men (Howard et al., 2013). Physiological and behavioral differences between black and white individuals may contribute to the observed difference in CRF. The oxidative capacity of muscle fibers is determined by their type. Type I muscle fibers are more oxidative and contain more mitochondria than type
II muscle fibers, which enables them to influence CRF to a greater extent than type II muscle fibers. Black individuals have been reported to have lower percent type I and higher percent type II muscle fiber proportions than white individuals (Ama et al., 1986, Suminski et al., 2002). Additionally, hemoglobin levels, which reflect the maximal amount of oxygen carrying capacity, have been reported to be lower in black than white individuals (Hunter et al., 2001, Beutler and West 2005, Swift et al., 2013a). However, adjustment for oxidative capacity and hemoglobin levels did not eliminate the difference in CRF (Hunter et al., 2001, Swift et al., 2013a), which suggests that these factors may contribute to the difference in CRF but do not explain it entirely.

A behavioral difference between black and white individuals that may contribute to the observed difference in CRF is engagement in physical activity. Recent data from the National Center for Health Statistics of the USA indicated that 57.4% of white adults met the minimum recommendations for leisure-time aerobic self-reported physical activity in 2018 compared to 46.4% of black adults (Public Tableau 2019). However, findings from studies in which physical activity was measured objectively using accelerometers are not consistent. In such studies, black individuals have been reported to be more sedentary (Diaz et al., 2017), less likely to meet the minimum recommendations for leisure-time aerobic physical activity (Lakoski and Kozlitina 2014), and less likely to engage in moderate-to-vigorous physical activity (Hooker et al., 2016) than white individuals. However, data obtained from the National Health and Nutritional Examination Survey (NHANES) 2003-2004 study cycle (National Center for Health Statistics 2019) showed similar levels of total physical activity and moderate-to-vigorous physical activity between black and white individuals (Hawkins et al., 2009). Additionally, Swift et al. reported similar pedometer-measured physical activity levels (5.087 steps/day and 4.930 steps/day) in black and white women (Swift et al., 2013a). Consequently, the extent to which the difference in CRF between black and white individuals can be explained by physical activity remains to be clarified. More studies using objective measurements of physical activity and CRF are needed.
A.5.6 Body fat

The relationship between body fat and CRF has been assessed in several studies using various methodologies and body fat indices. Studies in which CRF was objectively measured using a symptom-limited exercise test and body fat using BMI have consistently concluded that obese individuals have lower CRF than normal-weight individuals (Jette et al., 1990, So and Choi 2010, Lakoski et al., 2011, Stathokostas et al., 2015). Similarly, when body fat is measured using waist circumference and CRF is objectively measured using a symptom-limited exercise test there is a negative correlation between CRF and waist circumference that seems to be stronger in men than in women (Wedell-Neergaard et al., 2018, Dyrstad et al., 2019). Waist circumference has been reported to be more strongly correlated to CRF than BMI in a study in which the symptom-limited exercise test was terminated correctly at volitional exhaustion (Fogelholm et al., 2006). The opposite has been reported in another study (Dagan et al., 2013); however, in this study the symptom-limited exercise test was terminated when at least 85% of the predicted maximum heart rate was achieved, a termination criterion that should not be solely used to terminate an exercise test (Fletcher et al., 2013). Last, CRF has been found to be negatively correlated to body fat percentage (Schnurr et al., 2016) and the thickness of the epicardial fat tissue (Kim et al., 2010).

A.5.7 Socioeconomic status

The relationship between socioeconomic status (SES) and CRF has been assessed in several studies using various methodologies and SES indicators. A meta-analysis that included studies with large and representative samples (>300 participants) concluded that men and women with high education had higher CRF than men and women with low education (Hedges’ g effect size 0.12 and 0.19, respectively) (Ombrellaro et al., 2018). A study including 8471 participants also concluded that level of education and occupation correlated with SES even after multiple adjustments for demographic, clinical and socioeconomic variables (Shmueli et al., 2014). Living in a low SES neighborhood also seems to affect CRF. When CRF was estimated from submaximal
exercise testing, participants living in a low SES neighborhood exhibited 12% lower CRF than participants living in a high SES neighborhood (Lindgren et al., 2016). Similarly, when CRF was objectively measured using symptom-limited exercise stress testing, participants in the lowest tertile of neighborhood SES had 90% more chance to have very low CRF (lowest quintile) compared to participants living in a high SES neighborhood (Shishehbor et al., 2008). Recently, a study including 1.5 million men aged 18 – 19 years concluded that participants whose parents had low education had 44% - 119% more chance to have very low CRF (lowest tertile) compared to participants whose parents had high education (Lissner et al., 2019). The positive relationship between SES and CRF may be mediated by health-compromising behaviors that influence CRF, as SES has been reported to be associated with physical activity (Beenackers et al., 2012, Pudrovská and Anishkin 2013), adherence to dietary recommendations (Bonaccio et al., 2018, Lagstrom et al., 2019), obesity (McLaren 2007, Conklin et al., 2013, Shaikh et al., 2015), and smoking (Escobedo and Peddicord 1996, Hiscock et al., 2012). SES is also positively related to cardiovascular health (Foraker et al., 2019) and negatively related to the risk for coronary heart disease (Carlsson et al., 2016, Lee et al., 2017) and heart failure (Díaz-Toro et al., 2015), which are conditions that directly affect CRF and may therefore act as mediators to its relationship with SES.

A.6 Relationship between depression and cardiorespiratory fitness

It is evident from the abovementioned data that depression and CRF are both strong and independent predictors of mortality and disability, and that physical activity and exercise increase CRF and seem to reduce depression symptom severity. Given the potential interaction between depression and CRF through physical activity and exercise, it might be justifiable to assume that there is an inverse correlation between them. If such a correlation existed it would have prognostic and clinical implications, and it would stimulate further research on the effects of improving one variable on the other, as well as on the potential psychosomatic pathways linking depression and CRF. Elucidating such pathways would advance our understanding of the pathophysiology of depression, enhance our treatment strategies, and could provide further support to the
fact that depression is not only a psychiatric but also a biological disease with objectively measured adverse effects on biological indexes, such as CRF. This, in turn, might help alleviate the stigma associated to mental disease, which could result in more depressed patients asking and receiving help for their condition.

If such inverse correlation would be causative it would mean that a) depression could impair CRF and, respectively, b) treating depression could improve CRF; and/or c) improving CRF could reduce depressive symptoms and, respectively, d) worsening CRF could increase depressive symptoms. These effects might be mediated by biological and/or behavioral pathways. Physical activity can improve CRF (Garber et al., 2011) and has also been reported to enhance brain neuroplasticity (Hotting and Roder 2013) and mitigate neuroplasticity deficits in patients with depression (Phillips 2017). Inflammation, which is implicated in the development of depression (Howren et al., 2009, Valkanova et al., 2013), may also be attenuated by physical activity (Lavie et al., 2011, Mury et al., 2018). Last, elevated depressive symptoms may be demotivating in engaging with consistent physical activity and therefore, may lead to reduced CRF. Conversely, treating depression might enable individuals to become more physically active, which could lead to increased CRF.

It should be noted that a correlation between depression and CRF, causal or not, may be moderated and/or mediated by several factors that may affect the two variables. Physical activity is related to both depression and CRF, since it may be effective in reducing and preventing depressive symptoms (Kandola et al., 2019), while it maintains and improves CRF (Garber et al., 2011). Additionally, from a behavioral perspective, people with elevated depressive symptoms may not be motivated to engage in enough physical activity to maintain their CRF and therefore, may have impaired CRF. There is evidence that depression may be a significant risk factor for the development of sedentary lifestyle or decreased level of physical activity in otherwise healthy individuals (Roshanaei-Moghaddam et al., 2009) and in patients with heart disease (Sin et al., 2016). Determining whether physical activity acts as a mediator to a potential causative relationship between depression and CRF was one of the goals of our final study (see Chapter 6).
Sex, age, smoking history, BMI, socioeconomic status, and anxiety also influence either or both depression and CRF [see indicative references for sex (Laukkanen et al., 2009, Shanmugasegaram et al., 2012), age (Stordal et al., 2003, Laukkanen et al., 2009), smoking (Luger et al., 2014, Siddall et al., 2017), BMI (Laukkanen et al., 2009, Pereira-Miranda et al., 2017), socioeconomic status (Richardson et al., 2015, Ombrellaro et al., 2018), and anxiety (Pollack 2005, Loprinzi et al., 2017)] and therefore, may moderate and/or mediate a potential correlation between the two variables.

In order to assess whether a correlation between depression and CRF exists, we conducted a systematic review and meta-analysis of studies that reported an association between depression symptom severity and CRF in otherwise healthy adults, and of studies from which such an association could be calculated, regardless of the purpose of the study. This required contacting the authors of such studies and requesting the relevant correlation coefficient or the raw data so that we could calculate it ourselves.

B. Depressive symptoms and cardiorespiratory fitness in patients with heart disease

B.1 Epidemiology of depression in patients with heart disease

Heart disease patients provide an opportunity to understand the relationship between depression and CRF, and how this affects and is affected by their condition. In my professional position I work with heart disease patients in a hospital and therefore able to conduct original research in that setting.

Depression is quite common among patients with heart disease. There is extensive evidence estimating the rates of major depressive disorder in patients with coronary heart disease between 12% - 30% (Strik et al., 2001, Rudisch and Nemeroff 2003, Frasure-Smith and Lesperance 2006, Blumenthal 2008, Celano and Huffman 2011, Elderon and Whooley 2013), while the presence of elevated depressive symptoms may vary between 29% - 47% in this population (Ren et al., 2014, Doyle et al., 2015, Ghaemmobahamadi et al., 2018). The prevalence of depression in patients with heart failure has been reported to vary between 10% - 77.5% (Guck et al., 2003, Thomas et al., 2003, Yohannes et al., 2010, Hwang and Choi 2016) with difficulty to reach
consensus due to methodological weaknesses and the use of various diagnostic tools for depression (Yohannes et al., 2010). However, all these rates are at least two to three times higher than the rates of depression in the general population [4.4% (World Health Organization 2017)]. Respectively, patients with major depressive disorder have 56% - 80% higher chance of developing heart disease than the general population (Charlson et al., 2013, Penninx 2017).

B.2 The link between depression and heart disease

The high rates of comorbidity between depression and heart disease have stimulated extensive research on the psychobiological pathways that link these disorders, especially taking into account that they are the two of the three leading causes of global burden of disease (Dhar and Barton 2016). The linking pathways between depression and heart disease can be classified into three categories: behavioral, environmental, and biological.

- From a behavioral perspective, an acute cardiac event may trigger depression as an adjustment disorder (Hare et al., 2014) and respectively, depression may promote unhealthy lifestyle behaviors, including smoking, excessive alcohol use, physical inactivity, unhealthy diet (Penninx 2017), medication non-adherence (Seligman and Nemeroff 2015), and rehabilitation non-adherence (Pogosova et al., 2015), which may lead to the onset of heart disease or worsen the prognosis of existing heart disease.

- From an environmental perspective, depression seems to be predicted by neighborhood factors, including neighborhood socioeconomic conditions ([low socioeconomic status (SES)], more social security beneficiaries and more immigrants), physical factors (high levels of traffic noise) and social factors (lower social cohesion and less safety) (Generaal et al., 2019), although the evidence for traffic noise is of low quality (Dzhambov and Lercher 2019); however, according to a meta-analysis, neighborhood socioeconomic factors seem to have a transient (less than five years) effect on
depression (Richardson et al., 2015). There is also emerging evidence for the impact of air pollution on depression (Ali and Khoja 2019). Depression can also be predicted from childhood SES, including parental education (Gilman et al., 2002) and material hardship (Joinson et al., 2017). In a Spanish nationally representative sample adults who experienced a poor childhood financial situation were three times more likely to suffer from depression than those with a good childhood financial situation (Domenech-Abella et al., 2019). Parental (until the age of 15 years) and own occupational grade (Melchior et al., 2018), as well as own income (Kosidou et al., 2011), have also been reported to predict depression, while the occupational grade has been linked to depression trajectory, with lower grades predicting unfavorable depression trajectory and persistent depression (Melchior et al., 2013). Income and education, potentially mediated by financial strain and loneliness (Domenech-Abella et al., 2018), are significant predictors of mid-late life depression (Almeida et al., 2012) and their effect is similar in the UK and in China (Ruiz et al., 2019), which suggests that they are pervasive predictors of mid-late life depression in diverse social contexts. It seems that individual-level factors, including fewer material assets, lower education, female gender, economic inactivity and being divorced or widowed are more important determinants of depression than contextual factors, such as national income or country-level income inequality, as reported in a study that included participants from 53 countries (Rai et al., 2013).

Heart disease can also be predicted from environmental factors. Individuals who have low income and live in the most disadvantaged neighborhoods have three times the risk for developing CHD, compared with individuals who have high income and live in the most advantaged neighborhoods, after adjustment for established risk factors for CHD (Diez Roux et al., 2001), while the risk of myocardial infarction (Carlsson et al., 2016) and out-of-hospital fatal CHD and sudden fatal CHD (Foraker et al., 2011) is increased in residents from low-SES neighborhoods compared with residents in high-SES neighborhoods. Transportation noise (Vienneau et al., 2015) and traffic noise (Munzel et
al., 2017), as well as air pollution (Mustafic et al., 2012, Rajagopalan et al., 2018), have also been linked to CHD. CHD can be predicted from education level (Lee et al., 2017) and childhood and adult SES, which are derived from parental and own longest-held occupation, respectively (Lawlor et al., 2004). Cardiovascular death in women can also be predicted from income after adjustments for angiographic coronary disease, chest pain symptoms, and cardiac risk factors (Shaw et al., 2008). In a European study that included more than 95 million person-years cardiovascular mortality was 22% - 55% higher in men and 26% - 113% higher in women with lower educational level than in individuals with higher educational level, and this risk difference was higher in younger than older individuals (Avendano et al., 2006). Similarly, the incidence of heart failure is higher in residents of disadvantaged areas compared to residents of advantaged areas (Close et al., 2014, Cuthbertson et al., 2018) and when the heart failure is established the risk for all-cause hospitalization and all-cause mortality is higher in deprived patients than in affluent patients (Lawson et al., 2019). The precise mechanisms through which the above risks materialize remain elusive (Hawkins et al., 2012).

It is evident from the above findings that both depression and heart disease can be predicted from environmental factors, although the pathways through which these relationships evolve are unclear. Could depression and heart disease share some of those pathways or interact in longitudinal uni- or bio-directional relationships? More research is needed in this area.

- The bidirectional biological pathways between depression and heart disease seem more complex. The following dysfunctions may be present in patients with depression and patients with heart disease and may contribute to the development or worsening of depression and heart disease. These include:
  a) autonomic nervous system dysregulation, expressed by elevated levels of plasma catecholamines, elevated heart rate, low heart rate variability, exaggerated heart rate responses to physical stressors, high variability in ventricular repolarization, and low baroreceptor sensitivity (Carney et al., 2005).
b) immuno-inflammatory dysfunctions: Inflammation plays a significant role in the pathogenesis of heart disease and the prognosis of established heart disease. Coronary arterial inflammation is implicated in the development of atherosclerosis (Koenig 2001), plaque destabilization, and plaque rupture (Sarwar et al., 2009) forming a vascular inflammatory network that is composed of activated various leukocytes, vascular endothelial cells, vascular smooth muscle cells, platelets, excess reactive oxygen species, and cholesterol (Suzuki 2012). Chronic low-grade inflammatory activity has also been associated with structural heart disease (Kalogeropoulos et al., 2012), while myocarditis, regardless of its aetiology, can lead to impaired cardiac function, known as inflammatory cardiomyopathy (Cannie et al., 2019). When an acute coronary syndrome occurs the inflammation in response to ischemia and necrosis of cardiac tissue may initially facilitate tissue repair but subsequently it is involved in deleterious processes, including ventricular remodeling and myocardial fibrosis (Kalogeropoulos et al., 2012, Evora et al., 2013). Fibrinogen, C-reactive protein, lipoprotein-associated phospholipase A2, interleukin-6, serum amyloid A protein, serum albumin, and leucocyte count are among the most significant markers of inflammation that have been linked to heart disease (Danesh et al., 2000, Sarwar et al., 2009). There is evidence that immunoinflammatory abnormalities, including abnormal levels of inflammatory markers, such as interleukin-6, interleukin-1β, tumor necrosis factor α and interleukin-12, increased acute phase proteins, such as C-reactive protein, fibrinogen and haptoglobin, and abnormal complement factors, are not only involved in the pathogenesis of heart disease but also in the pathogenesis of depression (Maes et al., 2011, Nemeroff and Goldschmidt-Clermont 2012, Headrick et al., 2017, Liu et al., 2017). A “vicious circle” might be the following: Increased concentrations of the above proinflammatory cytokines accelerate atherosclerotic plaque progression and promote sickness behavior, which may lead to a physically inactive depressed lifestyle, which may further increase risk of heart disease.
c) endocrinal dysfunctions, including hyperactivity of the hypothalamus-pituitary adrenal axis (Kahl et al., 2018) and dysregulation of the endocrine factors insulin-like growth factor 1, fibroblast growth factor 21, and irisin (Wu et al., 2019) are implicated in the pathogenesis of both depression and heart disease.

d) increased susceptibility to blood clotting mainly due to platelet activation and aggregation: Patients with depression have been reported to have higher mean platelet volume – an indicator of platelet activity – than healthy individuals (Canan et al., 2012), while patients with heart disease and comorbid depression have been reported to exhibit increased platelet aggregation compared to patients with heart disease only (Williams et al., 2014, Seligman and Nemeroff 2015).

e) endothelial dysfunction: It is well established that endothelial dysfunction is present in atherosclerosis, which underlies heart disease, but it is also present in depression (Lippi et al., 2009, Celano and Huffman 2011), and it seems to be even worse in patients with heart disease and comorbid depression (Sherwood et al., 2005, Aydin Sunbul et al., 2017).

f) genomic factors: At least 24 genes that are involved in the pathogenesis of heart disease are also involved in the pathogenesis of depression (Amare et al., 2017). These genes are associated with heart disease risk factors, including hypertension, obesity, type II diabetes, and increased risk of arrhythmias and myocardial infarction, as well as with risk for depression and antidepressant treatment response (Kahl et al., 2018).

g) mitochondrial energy metabolism: It is well known that derangements in mitochondrial energy metabolism underpin the pathophysiology of heart failure, although it is not clear whether deranged mitochondrial energy metabolism contributes to pathological remodeling that leads to heart failure or it is an adaptive mechanism to an already diseased heart (Huss and Kelly 2005). Dysfunction in mitochondrial energy metabolism is also a consistent finding associated with depression (Zuccoli et al., 2017).
It should be noted that none of the above potentially bidirectional mechanisms that connect depression and heart disease has been reported to explain more than a small proportion of this relationship (Carney and Freedland 2017). More research is needed to clarify the nature and direction of the above pathways.

**B.3 Impact of depression on patients with heart disease**

Regardless of the types of mechanisms involved, the impact of depression on patients with heart disease is severe. Depression can serve as a barrier to behavior change and adoption of a healthy lifestyle and can worsen health-related quality of life (Dickens 2015, Pedersen et al., 2017). However, more importantly, depression has consistently been associated with increased all-cause and cardiovascular mortality risk in patients with CHD (de Miranda Azevedo et al., 2014, Berg et al., 2018, Zhang et al., 2018) that has been reported to be double (Leung et al., 2012) and even higher (Naqvi et al., 2005, Celano and Huffman 2011) compared to patients with CHD without depression. Taking this into account, the American Heart Association has elevated depression to the status of a risk factor for poor prognosis in patients with acute coronary syndrome (Lichtman et al., 2014). Similarly, depression in patients with heart failure increases the risk for rehospitalization, functional decline, and poorer quality of life (Freedland et al., 2011, Newhouse and Jiang 2014), and is an independent predictor of mortality (Norra et al., 2008, Nair et al., 2012, Celano et al., 2018) increasing the risk for all-cause mortality by 51% and for cardiovascular mortality by 119% (Fan et al., 2014).

Adding (and perhaps contributing) to the above is the fact that patients with CHD who also suffer from depression have a higher risk of non-adherence to prescribed medication and lifestyle modifications (Zellweger et al., 2004) estimated between 34% - 50%, compared to patients with CHD without depression (Kronish et al., 2006, Berg et al., 2018). Depressive symptoms in these patients have been reported to predict medication or lifestyle modifications non-adherence after three months (Rieckmann et al., 2006, Huffman et al., 2015), six months (Bauer et al., 2012), and even five years (Sin et al., 2016). Similar findings have been reported in patients with heart failure (Luyster et al., 2009, Alosco et al., 2013, Corotto et al., 2013).
Depression in patients with heart disease has also been associated with lower chance of attending a cardiac rehabilitation program (Kotseva et al., 2018) and even when these patients attend the program they are more likely to drop out of the program than patients without depression (Glazer et al., 2002, Casey et al., 2008, McGrady et al., 2009, Swardfager et al., 2011).

**B.4 Exercise and cardiorespiratory fitness in patients with heart disease**

Similarly to patients with depression only, physical activity and exercise as tools to treat depression have been studied in patients with depression together with established heart disease. There is evidence that physical activity (Janzon et al., 2015), exercise training (Lavie and Milani 2001, Lavie et al., 2011), and exercise-based cardiac rehabilitation (Jolly et al., 2006, Cao et al., 2016, Zheng et al., 2018) can reduce depressive symptoms in patients with heart disease, with a meta-analysis reporting an effect size of -0.38 (Tu et al., 2014), although these reductions might not be feasible in patients with minor depressive symptoms (Verschueren et al., 2018). Interestingly, depression in these patients has been reported to dampen the effect of exercise on CRF (Egger et al., 2008), implying the potential presence of biological pathways between depression and CRF.

The primary goal of exercise in patients with heart disease, though, is not the alleviation of depressive symptoms but the improvement of CRF. This is reflected in the expected outcomes of published cardiac rehabilitation guidelines (American Association of Cardiovascular and Pulmonary Rehabilitation 2013, Piepoli et al., 2014, Association of Chartered Physiotherapists in Cardiac Rehabilitation 2015) and is supported by the overwhelming evidence on the cardioprotective effects of enhanced CRF. Indeed, CRF is a strong and independent predictor of all-cause and cardiovascular mortality in patients with cardiovascular disease (Swift et al., 2013b, Barons et al., 2015, Harber et al., 2017), with each one MET increase being associated with 10% - 27% reduction in all-cause mortality (Martin et al., 2013, Artero et al., 2014, Taylor et al., 2016). Similarly, in patients with heart failure, each 5% improvement in CRF is associated with 10% reduction in all-cause mortality (Sabbag et al., 2018), while failure to improve CRF after exercise-based cardiac rehabilitation is
associated with double risk of all-cause mortality (Bakker et al., 2018). The cardioprotective effects of enhanced CRF are so strong that CRF has been found to attenuate the impact of dyslipidemia on cardiovascular mortality (Farrell et al., 2017) and the deleterious effects of obesity on cardiovascular health (Goel et al., 2011, McAuley et al., 2012, Lavie et al., 2013), potentially explaining the obesity paradox (McAuley and Beavers 2014, Piepoli et al., 2016), which is the inverse association between obesity and mortality that has been observed in patients with heart disease (Elagizi et al., 2018).

**B.5 Relationship between depression and cardiorespiratory fitness in patients with heart disease**

It is evident from the abovementioned data that depression and CRF are both strong and independent predictors of mortality and disability in patients with heart disease, and that exercise increases CRF and seems to reduce depression symptom severity in this population. Given the beneficial effects of exercise on depression and CRF it might be justifiable to assume that there is an inverse correlation between depression and CRF in patients with heart disease. If this correlation existed it would provide important information that could help clinicians in risk stratification, patient assessment, goal setting, treatment, and rehabilitation. Moreover, it would encourage further research on the effects of treating one variable on the other, which might lead to enhanced therapeutic and rehabilitation services. Finally, it might suggest the presence of psychosomatic pathways between depression and CRF in patients with heart disease, the elucidation of which would advance our understanding of the pathophysiology of depression in this population.

In order to assess whether there is a correlation between depression and CRF in patients with heart disease, we conducted a systematic review and meta-analysis of studies reporting an association between depression symptom severity and CRF in patients with heart disease and of studies from which such an association could be calculated, regardless of the purpose of the study. Similarly to our meta-analysis in otherwise healthy individuals, this required contacting the authors of those studies and requesting the relevant correlation coefficient or the raw data so that we could calculate it ourselves.
C. Moderators of the relationship between depression and CRF in patients with heart disease

As mentioned before, a potential relationship between depression and heart disease may be moderated by several factors, including physical activity, sex, age, smoking history, BMI, socioeconomic status, and anxiety. In both our meta-analyses we were not able to assess the effect of such moderators at within-study level because raw data were not available for every study. Assessing potential moderators of this relationship would be important because it would inform clinicians of specific characteristics of patients that would strengthen or weaken the relationship and therefore help clinicians take better informed decisions regarding patient management. In patients with heart disease potential moderators of the relationship between depression and CRF might be sex, age, BMI, cardiovascular diagnosis, presence of diabetes, smoking history, socioeconomic status, and anxiety because they are related with either or both depression and CRF [see indicative references for sex (Laukkanen et al., 2009, Shanmugasegaram et al., 2012), age (Stordal et al., 2003, Laukkanen et al., 2009), BMI (Laukkanen et al., 2009, Pereira-Miranda et al., 2017), cardiovascular diagnosis (Laukkanen et al., 2009), presence of diabetes (Moulton et al., 2015, Zaccardi et al., 2015), smoking history (Luger et al., 2014, Siddall et al., 2017), socioeconomic status (Lorant et al., 2003, Ombrellaro et al., 2018), and anxiety (Pollack 2005, Loprinzi et al., 2017)]. In order to conduct a moderation analysis in the relationship between depression and CRF in patients with heart disease I would need a large sample size, which was impossible to have in my workplace; I therefore used the dataset from an Open Access study (Barons et al., 2015), which included a large sample of patients with heart disease (n = 2714).

D. Translation and validation of the Cardiac Depression Scale to Arabic

After the completion of the two abovementioned systematic reviews and meta-analyses and the moderation analysis on the correlation between depression and CRF the next step was to assess whether there is causation between depression and CRF in patients with heart disease; that is whether depression reduces CRF in patients with heart disease. This would have required assessment of depression symptom severity in
patients with heart disease and therefore, a depression rating scale would have been needed. There are several psychometric scales that have been used to assess depression in patients with heart disease, including the Beck Depression Inventory II (Beck et al., 2010), the Hospital Anxiety and Depression Scale (Zigmond and Snaith 1983), the Hamilton Rating Scale for Depression (Hamilton 1960), and the Centre for Epidemiologic Studies Depression Scale (Radloff 1977). Although these scales are well validated, they have not been developed specifically for patients with heart disease and they have not been validated in this population. As a result, these scales may not be comprehensive or sensitive enough to assess all aspects of depressive symptomatology, including somatic symptoms, or lesser depressive symptoms that are nevertheless clinically significant in patients with heart disease (Bush et al., 2001, Catipovic-Veselica et al., 2007). The Cardiac Depression Scale (CDS) (Hare and Davis 1996), on the other hand, is the only psychometric scale that was specifically developed to measure depressive symptoms in patients with heart disease and has been validated several times in this population (Hare and Davis 1996, Birks et al., 2004, Wise et al., 2006, Kiropoulos et al., 2012, Ski et al., 2012). We therefore decided to measure depression symptom severity using the CDS in the study assessing whether depression reduces CRF in patients with heart disease.

This study would have been conducted in my country of residence, Qatar, which is an Arabic country with the majority of the population being Arabs and having the Arabic language as mother tongue. Therefore, in order to ensure that the measurement of depression symptom severity would not be biased, the Arabic version of the CDS would have needed to be used for the majority of the participants. Since there was no Arabic version of the CDS, we conducted a translation and validation study of the CDS in this population.

The prevalence of depression in the general population of Qatar ranges from 4.2% to 6.6% (Khaled 2019) and it reaches up to 18.3% in patients who present to a primary care physician (Bener et al., 2015). In patients with heart disease the prevalence of mild mood disturbance and clinical depression has been reported to be 15% and 5%, respectively, with female patients being at double risk than male patients (Donnelly et al., 2016), although these estimations should be interpreted with caution because this study used a non-probability
convenient sampling technique to recruit participants. Cardiovascular disease is the second cause of death (after injuries) in Qatar (Chaabna et al., 2018), while 83% of the population is engaged in low or no physical activity (Qatar Biobank 2019). Considering the prevalence of depression, the impact of cardiovascular disease, and the high inactivity levels in the population of Qatar, research that might reveal a causal relationship between depression and CRF in patients with heart disease would be of interest.

**E. Is the correlation between depression and cardiorespiratory fitness causative?**

Having conducted two meta analyses on the correlation between depression and CRF, a moderation analysis of this correlation, and a translation and validation of the CDS in Arabic, the next step was to assess whether the correlation between depression and CRF was causative. However, due to delays that are detailed in the sixth chapter of this thesis and the time limits of the PhD, this study was regrettably abandoned for PhD purposes.

Consequently, this thesis consists of the following four studies and one research proposal:

1. Is there an association between depression symptom severity and CRF in otherwise healthy individuals? (Papasavvas et al., 2016b)

2. Is there an association between depression symptom severity and CRF in patients with heart disease? (Papasavvas et al., 2017)


4. Translation and validation of the Cardiac Depression Scale to Arabic. (Papasavvas et al., 2016a)

5. Does depression reduce CRF? A Quasi-Experimental Study (research proposal)
References


Williams, M. D., Harris, R., Dayan, C. M., Evans, J., Gallacher, J., Ben-Shlomo, Y. 2009. Thyroid function and the natural history of depression: findings from the Caerphilly Prospective Study (CaPS) and a meta-analysis. Clin Endocrinol (Oxf), 70(3), 484-92.


Chapter 2

Is there an association between depression symptom severity and CRF in otherwise healthy individuals?
Depression symptom severity and cardiorespiratory fitness in healthy and depressed adults: a systematic review and meta-analysis

Theodoros Papasavvas¹ ² • Robert O. Bonow³ • Mohammad Alhashemi¹ • Dominic Micklewright²

Published online: 7 October 2015
© Springer International Publishing Switzerland 2015

Abstract

Background: Depression symptom severity and cardiorespiratory fitness (CRF) are significant predictors of mortality and disability. However, the relationship between the two is unclear.

Objective: This meta-analysis assessed the relationship between depression symptom severity and CRF in healthy and depressed adults (aged 18 years and over).

Search Methods: The PubMed, Cochrane Library, Google Scholar and ProQuest databases were browsed for relevant English-language studies published from January 2000 to August 2014.

Selection Criteria: Studies reporting a correlation between a depression scale and maximum oxygen consumption (VO₂peak), as well as studies from the data of which such a correlation could be calculated, were included in this analysis.

Data Analysis: Correlation coefficients (CCs) were converted to Fisher’s z values, and the analysis was performed using a random-effects model. Then, summary effects and 95 % confidence intervals (CIs) were converted back to CCs.

Results: Sixteen studies (totaling 4039 participants) were included in this analysis. A modest correlation between depression symptom severity and CRF was found (CC -0.16, 95 % CI -0.21 to -0.10), appearing

1 Cardiac Rehabilitation Department, Heart Hospital, Hamad Medical Corporation, Doha, Qatar
2 School of Biological Sciences, University of Essex, Colchester, UK
3 Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA
stronger in male participants (CC - 0.22, 95 % CI -0.26 to -0.18) than in female participants (CC -0.12, 95 % CI -0.19 to -0.05; p = 0.01). There was no difference in the summary effect between healthy and depressed adults (p = 0.43). Heterogeneity was moderate (I² = 33 %; p = 0.09).

Conclusions: Depression symptom severity is inversely correlated with CRF, and this correlation is stronger in men than in women. Clinical and prognostic implications of the correlation are discussed. These findings should stimulate further research on the effects of treating one variable on the other.

<table>
<thead>
<tr>
<th>Key points</th>
</tr>
</thead>
<tbody>
<tr>
<td>In this meta-analysis, a modest inverse correlation between depression symptom severity and cardiorespiratory fitness was found, appearing stronger in men than in women.</td>
</tr>
<tr>
<td>This correlation has clinical and prognostic implications and should stimulate further research on the effects of treating one variable on the other.</td>
</tr>
</tbody>
</table>

1 Introduction

Depression is the leading cause of disability worldwide, afflicting more than 350 million people of all ages [2]. It is also associated with development of an array of diseases, including coronary artery disease [3], type 2 diabetes mellitus (DM) [4] and dementia [5], as well as musculoskeletal disorders [6]. Furthermore, comorbid depression is associated with a poor prognosis of the underlying disease, such as DM [7] or heart failure [8], while also increasing the mortality risk after myocardial infarction [9]. Adding to this, non-adherence to medication is common in patients with depression alone [10] and in patients with chronic diseases and comorbid depression [11], imposing an additional inherent difficulty in patient management.

The role of exercise as a treatment modality for depression has been extensively studied. Exercise seems to reduce depression symptom severity [12], and several meta-analyses have reported therapeutic effects of physical activity (PA) interventions in patients with clinical depression, with effects sizes ranging from 0.34 to 0.8, in comparison with control or placebo groups [13–16], although effect sizes may be smaller in
methodologically robust trials [14]. Reduced depression symptom severity has also been reported in patients with chronic diseases [17,18], as well as in healthy individuals [19].

It is well known that regular exercise increases cardiorespiratory fitness (CRF) [20], which is a strong predictor of all-cause mortality in healthy individuals [21, 22] and an established prognostic factor for cardiovascular disease [23, 24]. Given the beneficial effects of exercise on depression and the effect of regular exercise on CRF, it could be argued that there may be an inverse correlation between depression symptom severity and CRF.

If such a correlation exists, it would support the rationale for improving CRF as a means to treat depression and vice versa. Furthermore, if potential moderators, such as PA, are controlled, this correlation would imply the presence of inherent biological pathways between depression and CRF, linking mood to objectively measured biological indices. Studying these pathways might advance our understanding of the biological pathophysiology of depression. Such a correlation would also link depression symptom severity to mortality, which could have an impact on clinical practice.

The correlation between depression symptom severity and CRF has been reported in few studies [25–30], of which only two [26, 28] used a symptom-limited cardiopulmonary exercise test (CPET) to measure maximum oxygen consumption (VO_{2max}) in order to assess CRF. VO_{2max} is considered the best index of CRF [31], and indirect estimations of CRF (e.g. from the time and workload of the exercise test) could bias the correlation between depression symptom severity and CRF. Although studies reporting this correlation are few, this correlation can be assessed in every study that reports depression symptom severity and VO_{2max}, regardless of whether this is the purpose of the study or not. We therefore conducted a meta-analysis of cross-sectional correlations between depression symptom severity and CRF, which were calculated from studies that reported depression symptom severity and VO_{2max}.
2 Methods

2.1 Eligibility Criteria

Studies published between 2000 and 2014 were inspected. The restriction to recent studies was chosen on the rationale that access to the raw databases of the studies in order to calculate correlations would be needed and the availability of such databases would diminish with time. Experimental and observational studies in adult healthy participants and patients with major depressive disorder (MDD), published and unpublished in English, were eligible for inclusion. MDD was defined by using the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) [32] or the International Classification of Diseases, Tenth Revision (ICD-10) [33], or by scoring the above defined cut-off scores in validated scales. Studies in patients with MDD were eligible in order for the total range of depressive symptoms to be used for calculation of the correlation between depression symptom severity and CRF. Studies that reported correlation(s) between CRF and depression symptom severity, as well as studies from which such correlations could be calculated, were included. Maximum oxygen consumption (VO$_{2\text{max}}$ or VO$_{2\text{peak}}$), measured using a symptom-limited CPET, was used for measuring CRF, and validated scales measuring depressive symptoms were used for measuring depression symptom severity.

In order to reduce measurement error, studies with estimated VO$_{2\text{max}}$ or VO$_{2\text{peak}}$ values (from watts, speed, time, etc.) were excluded. Studies including obese participants, pregnant or postnatal women, and participants in smoking abstinence were also excluded because these conditions have been linked to depression and therefore may act as additional moderators of the relationship between depression symptom severity and CRF, the study of which was out of the scope of this meta-analysis.
2.2 Information Sources

A systematic computerized literature search in the databases of PubMed, Cochrane Library, Google Scholar and ProQuest, covering articles from January 2000 to August 2014, was conducted. ProQuest contains unpublished theses, inclusion of which reduces publication bias. Reference lists from the publications of the included studies were also reviewed, and the principal investigators of the included studies were asked to provide relevant studies.

2.3 Searches

The titles and abstracts (PubMed), titles (Google Scholar), abstracts (ProQuest) and full text (Cochrane Library) of study publications were searched using the following search terms: exercise AND depress*, stress test AND depress*, cardiorespiratory AND depress*, aerobic capacity AND depress*, fitness AND depress*, oxygen uptake AND depress*, VO\textsubscript{2max} AND depress*, VO\textsubscript{2peak} AND depress*.

All search terms were followed by NOT menopause NOT pregnant NOT postpartum NOT postnatal NOT fibromyalgia NOT Parkinson’s NOT stroke NOT sclerosis NOT lumbar NOT spondylitis NOT osteoarthritis NOT fracture NOT musculoskeletal NOT cancer NOT bipolar NOT binge NOT adolescent* NOT rat, except in the Google Scholar search.

2.4 Study Selection and Data Collection Process

All potentially relevant study publications, judged from their titles and abstracts, were acquired in full-text form, or their authors were contacted and asked if VO\textsubscript{2max} or VO\textsubscript{2peak} values and a depression scale score had been reported. Studies reporting VO\textsubscript{2max} or VO\textsubscript{2peak} values and a depression scale score, or a correlation between them, were included in the meta-analysis. Authors of studies reporting VO\textsubscript{2max} or VO\textsubscript{2peak} values and a depression scale score were contacted, and the raw data or the correlation were requested.
2.5 Data Items

The correlation between CRF and depression symptom severity was the only data item for this meta-analysis and was either reported or calculated from the included studies. In cases of prospective or trial studies, the correlation between CRF and depression symptom severity at baseline was reported or calculated. In cases of studies with more than one group, all groups were merged and the correlation between CRF and depression symptom severity was calculated using the raw data from the total cohort.

2.6 Risk of Bias in Individual Studies

In order to eliminate the risk of selective outcome reporting and to reduce the risk of selection bias in prospective or trial studies, the correlation between CRF and depression symptom severity at baseline only was reported or calculated; therefore, other types of bias—such as performance bias, detection (or ascertainment) bias, attrition bias and reporting bias—were not assessed, because they do not affect baseline measurements. The risk of selection bias was assessed by reporting blinding of participants and data collectors to allocation because, although the correlation was calculated for the total cohort, participants’ and data collectors’ awareness of the allocation could influence performance and measurement, respectively. Additionally, the risk of measurement error for CRF was assessed by reporting use of a peak respiratory exchange ratio (RER) of ≥1.1 and use of a 6–20 Borg Scale of perceived exertion [34] rating of >18 as a termination criterion for the CPET. The former is considered an objective indicator of sufficient subject effort [35], and the latter is considered to suggest maximal effort during a CPET [34]. The risk of measurement error for depression symptom severity was assessed by reporting whether the depression scale scoring was researcher completed, because self-administered scale scoring might induce bias.
2.7 Summary Measures

Pearson’s (or Spearman’s, in the case of non-parametric data) correlation coefficients (CCs) between VO$_{2\text{max}}$ or VO$_{2\text{peak}}$ and the depression scale score were used to create the summary measure. A random-effects model was used, since the effect of moderators (e.g. sex, age) on the true effect size was unknown.

2.8 Synthesis of Results

The method of Hedges and Olkin [36] was used for synthesizing the CCs. In this method, the correlations are converted to the Fisher’s z scale [37] and the analysis is performed using the transformed values. Then, the results, such as the summary effects and 95 % confidence intervals (CIs), are converted back to correlations. This method of converting correlations to Fisher’s z values was chosen in order to eliminate bias in the correlation variance, since it depends strongly on the correlation itself, as can be seen in the following formula [38]:

\[ V_r = \frac{(1 - r^2)^2}{n - 1} \]

where $V_r$ is the variance of $r$, $r$ is the correlation coefficient and $n$ is the sample size.

Some inconsistency in the results of studies (within-study variation) is expected because of sampling error. Any excessive inconsistency (between-studies variation) is considered to reflect true differences between studies and is called heterogeneity. Heterogeneity was assessed using Cochran’s Q $\chi^2$ test and using $I^2$, which represents the percentage of the total variation that is attributed to true differences between studies. Confidence (uncertainty) intervals for $I^2$ values were also calculated.

MIX 2.0 software [39] and comprehensive meta-analysis software (Englewood, NJ, USA) were used for statistical analysis.
2.9 Risk of Bias Across Studies

In 15 of 16 included studies, the effect sizes were not reported; instead, the authors were contacted and asked to provide raw data. Therefore, the risk of publication bias was expected to be minor. Nevertheless, publication bias was assessed visually on a funnel plot (the effect size by the inverse of its standard error) and statistically using Egger’s test, the Begg and Mazumdar rank correlation test and the trim-and-fill method [40]. It should be mentioned, however, that other factors, such as study quality and true heterogeneity, can produce asymmetry in funnel plots. Duplication was avoided by juxtaposing the authors’ names, the study’s sample size, and the subjects’ sex and age. In cases of ambiguity, the authors were contacted. Availability bias (selective inclusion of studies that are easily accessible to the researcher) with regard to omitting studies from which the required data were difficult to acquire was addressed by contacting all authors of the study and by regular (biweekly) follow-ups.

2.10 Additional Analyses

Prespecified exclusion sensitivity analysis was performed to assess the effect of exclusion of each study (one at a time) on the summary effect. Three post hoc subgroup analyses were performed, using a Q test based on analysis of variance [38] for a random-effects model with a pooled estimate of \( \tau^2 \) (between-studies dispersion). The pooled estimate of \( \tau^2 \) was used because there was no reason to assume different true study-to-study dispersion between subgroups and because if there was indeed different true study-to-study dispersion between subgroups, a sufficient number of studies within each subgroup would be needed to yield an accurate estimate of it, which was not the case in this meta-analysis.

In the first subgroup analysis, the moderator was ‘sex’ with two levels: male and female. In the second subgroup analysis, the moderator was ‘participants suffered from MDD’ with two levels: yes and no. In the third subgroup analysis, the moderator was ‘RER \( \geq 1.1 \) was a criterion for subject’s sufficient effort during a
CPET with two levels: yes and no. Prespecified meta-regression was performed to assess the effect of age and publication year on the effect sizes.

3 Results

3.1 Study Selection

Study selection was conducted by T.P and resulted in 16 studies to be included in the analysis. The search in the databases provided a total of 7768 citations. Of these, 7491 were discarded after review of the title and the abstract, since it was clear that they did not contain assessment of CRF and depression symptom severity. After acquisition of the full text and/or contact with the authors of the remaining studies, 211 studies were discarded because they did not contain assessment of CRF and/or depression symptom severity, while 41 studies were discarded because CRF was assessed without direct measurement of $\text{VO}_{2\text{max}}$ or $\text{VO}_{2\text{peak}}$—instead, CRF was estimated from watts, speed, time, etc. Of the remaining 25 eligible studies, four were duplicates, while five studies were excluded because it was impossible to acquire the relevant data from their authors. See the flow diagram in Fig. 2.1.

3.2 Study Characteristics

Sixteen studies were included in the analysis, totaling 4039 participants (see Table 2.1). One cross-sectional study [28] assessed and reported the correlation between depression and $\text{VO}_{2\text{peak}}$. The purpose of the 15 remaining studies [41–55] was different; therefore, their types varied, including cross-sectional, longitudinal and trial studies. Raw data were provided in eight studies and CCs were provided in seven studies. One CC per study was calculated and included in the analysis. If a study used a depression scale that was completed by a clinician (e.g. the Hamilton Depression Rating Scale) and a self-completed depression scale (e.g. the Beck Depression Inventory) the former was used for the calculation of the correlation, since it has been suggested that observer-rated scales may be more valid than self-rated scales [56]. If a study had more than one group,
the correlation for the total cohort was calculated. This included mixing of depressed participants with healthy controls, and it was performed in order for the maximum range of the depression scale scores to be used for calculation of the correlation. If a study had more than one measurement, the correlation at baseline was calculated. Pearson’s $r$ values and Spearman’s $\rho$ values were calculated in the case of parametric and non-parametric data, respectively.

Flow diagram of study selection. CRF cardiorespiratory fitness, $VO_{2\text{max}}$ maximum oxygen consumption, $VO_{2\text{peak}}$ maximum oxygen consumption
<table>
<thead>
<tr>
<th>References</th>
<th>n</th>
<th>Age [years; mean]</th>
<th>Sex</th>
<th>Study type</th>
<th>Scale</th>
<th>Mode of exercise test</th>
<th>Presence of MDD</th>
<th>Details for the correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hollenberg et al. [46]</td>
<td>634</td>
<td>66.5</td>
<td>F</td>
<td>Longitudinal</td>
<td>CES-D</td>
<td>Treadmill</td>
<td>Both healthy and MDD participants</td>
<td>Total cohort, Spearman's ρ</td>
</tr>
<tr>
<td>Stewart et al. [54]</td>
<td>53</td>
<td>62.2</td>
<td>M</td>
<td>Cross-sectional</td>
<td>POMS-D</td>
<td>Treadmill</td>
<td>Both healthy and MDD participants</td>
<td>Spearman's ρ, females excluded (BMI &gt; 30 kg/m²)</td>
</tr>
<tr>
<td>Kiive et al. [48]</td>
<td>46</td>
<td>43.5</td>
<td>M</td>
<td>Trial</td>
<td>MADRS</td>
<td>Cycle ergometer</td>
<td>Both healthy and MDD participants</td>
<td>Total cohort, baseline</td>
</tr>
<tr>
<td>St Clair Gibson et al. [53]</td>
<td>9</td>
<td>38.1</td>
<td>M, F</td>
<td>Trial</td>
<td>BDI-SF</td>
<td>Treadmill</td>
<td>Only healthy participants</td>
<td>Only control group; acquired training intolerance group excluded</td>
</tr>
<tr>
<td>Hoffman et al. [45]</td>
<td>200</td>
<td>51.7</td>
<td>M, F</td>
<td>RCT</td>
<td>HAM-D</td>
<td>Treadmill</td>
<td>Only participants with MDD</td>
<td>Total cohort, baseline</td>
</tr>
<tr>
<td>Boettger et al. [42]</td>
<td>44</td>
<td>37</td>
<td>M, F</td>
<td>Trial</td>
<td>HAM-D</td>
<td>Cycle ergometer</td>
<td>Both healthy and MDD participants</td>
<td>Total cohort, baseline</td>
</tr>
<tr>
<td>Chu et al. [42]</td>
<td>39</td>
<td>26.4</td>
<td>F</td>
<td>RCT</td>
<td>BDI</td>
<td>Treadmill</td>
<td>Only participants with MDD</td>
<td>Total cohort, baseline</td>
</tr>
<tr>
<td>Krogh et al. [50]</td>
<td>165</td>
<td>38.9</td>
<td>M, F</td>
<td>RCT</td>
<td>HAM-D</td>
<td>Cycle ergometer</td>
<td>Only participants with MDD</td>
<td>Total cohort, baseline</td>
</tr>
<tr>
<td>Valtonen et al. [55]</td>
<td>242</td>
<td>53</td>
<td>M</td>
<td>Cross-sectional</td>
<td>HPL</td>
<td>Cycle ergometer</td>
<td>Both healthy and MDD participants</td>
<td>Total cohort</td>
</tr>
<tr>
<td>Donath et al. [44]</td>
<td>30</td>
<td>36</td>
<td>F</td>
<td>Trial</td>
<td>HAM-D</td>
<td>Cycle ergometer</td>
<td>Both healthy and MDD participants</td>
<td>Total cohort</td>
</tr>
<tr>
<td>Rice et al. [28]</td>
<td>131</td>
<td>66.3</td>
<td>M, F</td>
<td>Cross-sectional</td>
<td>BDI</td>
<td>Treadmill</td>
<td>Only healthy participants</td>
<td>Step 1 for both genders</td>
</tr>
<tr>
<td>Krogh et al. [51]</td>
<td>115</td>
<td>41.6</td>
<td>M, F</td>
<td>RCT</td>
<td>HAM-D</td>
<td>Cycle ergometer</td>
<td>Only participants with MDD</td>
<td>Total cohort, baseline</td>
</tr>
<tr>
<td>Raso et al. [52]</td>
<td>37</td>
<td>67.9</td>
<td>F</td>
<td>Cross-sectional</td>
<td>POMS-D</td>
<td>Treadmill</td>
<td>Both healthy and MDD participants</td>
<td>Total cohort, baseline</td>
</tr>
<tr>
<td>Dery et al. [43]</td>
<td>24</td>
<td>22</td>
<td>M, F</td>
<td>Trial</td>
<td>BDI</td>
<td>Cycle ergometer</td>
<td>Only healthy participants</td>
<td>Total cohort, baseline</td>
</tr>
<tr>
<td>References</td>
<td>n</td>
<td>Age [years; mean]</td>
<td>Sex</td>
<td>Study type</td>
<td>Scale</td>
<td>Mode of exercise test</td>
<td>Presence of MDD</td>
<td>Details for the correlation</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>----</td>
<td>-------------------</td>
<td>-----</td>
<td>------------</td>
<td>-------------</td>
<td>-----------------------</td>
<td>----------------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>Hunter et al. [47]</td>
<td>25</td>
<td>66</td>
<td>F</td>
<td>RCT</td>
<td>POMS</td>
<td>Treadmill</td>
<td>Only healthy participants</td>
<td>Total cohort, baseline</td>
</tr>
<tr>
<td>Kreuzfeld et al. [49]</td>
<td>59</td>
<td>54</td>
<td>M, F</td>
<td>Trial</td>
<td>BDI</td>
<td>Cycle ergometer</td>
<td>Both healthy and MDD participants</td>
<td>Total cohort, baseline</td>
</tr>
</tbody>
</table>

*BDI Beck Depression Inventory, BDI-SF Beck Depression Inventory—Short Form, BMI body mass index, CES-D Center for Epidemiologic Studies—Depression Scale, F female, HAM-D Hamilton Depression Rating Scale, HPL Human Population Laboratory Depression Scale, M male, MARDS Montgomery–Asberg Depression Rating Scale, MDD major depressive disorder, POMS-D Profile of Mood States—Depression Subscale, RCT randomized controlled trial*
3.3 Risk of Bias Within Studies

The results of the assessment of risks of selection bias and measurement errors can be seen in Table 2.2.

<table>
<thead>
<tr>
<th>References</th>
<th>Participants blinded to allocation</th>
<th>Data collectors blinded to allocation</th>
<th>RER $\geq$ 1.1 used</th>
<th>RPE &gt; 18 used</th>
<th>Researcher-completed depression scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hollenberg et al. [46]</td>
<td>N/A</td>
<td>N/A</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Stewart et al. [54]</td>
<td>N/A</td>
<td>N/A</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Kiive et al. [48]</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>St Clair Gibson et al. [53]</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Hoffman et al. [45]</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Boettger et al. [41]</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Chu et al. [42]</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Krog et al. [50]</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Valtonen et al. [55]</td>
<td>N/A</td>
<td>N/A</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Donath et al. [44]</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Rice et al. [28]</td>
<td>N/A</td>
<td>N/A</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Krog et al. [51]</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Raso et al. [52]</td>
<td>N/A</td>
<td>N/A</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Dery et al. [43]</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Hunter et al. [47]</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Kreuzfeld et al. [49]</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Risk of bias within studies. *N/A* not applicable, *RER* respiratory exchange ratio, *RPE* rates of perceived exertion

3.4 Results of Individual Studies

A report of the effect sizes, including sample sizes, 95% CIs, $z$ statistics and study weights, can be seen in Table 2.3.
### Table 2.4

<table>
<thead>
<tr>
<th>References</th>
<th>n</th>
<th>CC</th>
<th>95% CI</th>
<th>z value</th>
<th>P value</th>
<th>Weight [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hollenberg et al. 2003 [46]</td>
<td>634</td>
<td>-0.11</td>
<td>-0.19 to -0.03</td>
<td>-2.77</td>
<td>0.01</td>
<td>17.67</td>
</tr>
<tr>
<td>Stewart et al. 2003 [54]</td>
<td>53</td>
<td>-0.25</td>
<td>-0.49 to 0.02</td>
<td>-1.81</td>
<td>0.07</td>
<td>3.63</td>
</tr>
<tr>
<td>Kiive et al. 2004 [48]</td>
<td>46</td>
<td>-0.29</td>
<td>-0.54 to 0.00</td>
<td>-1.96</td>
<td>0.05</td>
<td>3.18</td>
</tr>
<tr>
<td>St Clair Gibson et al. 2006 [53]</td>
<td>9</td>
<td>-0.11</td>
<td>-0.72 to 0.60</td>
<td>-0.27</td>
<td>0.79</td>
<td>0.49</td>
</tr>
<tr>
<td>Hoffman et al. 2008 [45]</td>
<td>200</td>
<td>0.01</td>
<td>-0.13 to 0.15</td>
<td>0.14</td>
<td>0.89</td>
<td>10.19</td>
</tr>
<tr>
<td>Boettger et al. 2009 [41]</td>
<td>44</td>
<td>-0.38</td>
<td>-0.61 to -0.09</td>
<td>-2.56</td>
<td>0.01</td>
<td>3.05</td>
</tr>
<tr>
<td>Chu et al. 2009 [42]</td>
<td>39</td>
<td>-0.18</td>
<td>-0.47 to 0.14</td>
<td>-1.09</td>
<td>0.27</td>
<td>2.71</td>
</tr>
<tr>
<td>Krogh et al. 2009 [50]</td>
<td>165</td>
<td>-0.15</td>
<td>-0.30 to 0.00</td>
<td>-1.92</td>
<td>0.05</td>
<td>8.99</td>
</tr>
<tr>
<td>Valtonen et al. 2009 [55]</td>
<td>2428</td>
<td>-0.22</td>
<td>-0.26 to -0.18</td>
<td>-11.01</td>
<td>&lt; 0.001</td>
<td>23.45</td>
</tr>
<tr>
<td>Donath et al. 2010 [44]</td>
<td>30</td>
<td>-0.36</td>
<td>-0.64 to 0.00</td>
<td>-1.96</td>
<td>0.05</td>
<td>2.09</td>
</tr>
<tr>
<td>Rice et al. 2010 [28]</td>
<td>131</td>
<td>-0.20</td>
<td>-0.36 to -0.03</td>
<td>-2.29</td>
<td>0.02</td>
<td>7.65</td>
</tr>
<tr>
<td>Krogh et al. 2012 [51]</td>
<td>115</td>
<td>-0.15</td>
<td>-0.32 to 0.03</td>
<td>-1.60</td>
<td>0.11</td>
<td>6.94</td>
</tr>
<tr>
<td>Raso et al. 2012 [52]</td>
<td>37</td>
<td>-0.05</td>
<td>-0.37 to 0.28</td>
<td>-0.29</td>
<td>0.77</td>
<td>2.58</td>
</tr>
<tr>
<td>Dery et al. 2013 [43]</td>
<td>24</td>
<td>-0.03</td>
<td>-0.43 to 0.38</td>
<td>-0.14</td>
<td>0.89</td>
<td>1.65</td>
</tr>
<tr>
<td>Hunter et al. 2013 [47]</td>
<td>25</td>
<td>-0.05</td>
<td>-0.44 to 0.35</td>
<td>-0.23</td>
<td>0.81</td>
<td>1.73</td>
</tr>
<tr>
<td>Kreuzfeld et al. 2013 [49]</td>
<td>59</td>
<td>0.02</td>
<td>-0.24 to 0.27</td>
<td>0.15</td>
<td>0.88</td>
<td>4.00</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>4039</td>
<td>-0.16</td>
<td>-0.21 to -0.10</td>
<td>-5.41</td>
<td>&lt; 0.001</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Details of all effect sizes. CC correlation coefficient, CI 95% confidence interval.

### 3.5 Synthesis of Results

The summary CC was -0.16 with a 95% CI of -0.21 to -0.10 (see Fig. 2). Although heterogeneity was not statistically significant, it was considerable ($I^2 = 33.6\%$, 95% CI 0–63.6%; $Q = 22.3$; degrees of freedom [df] = 15; $p = 0.09$). Because of the presence of considerable heterogeneity and the fact that the power for the Cochran’s $\chi^2$ test was low because of the small number of included studies, it was decided to explore the heterogeneity.
**Figure 2.2**

Forest plot of effect sizes of correlation between depression symptom severity and cardiorespiratory fitness. CC correlation coefficient, CI confidence interval.
3.6 Risk of Bias Across Studies

The visual impression of the funnel plot (the effect size by the inverse of its standard error) showed no sign of publication bias (see Electronic Supplementary Material Fig. S2.1). This was confirmed statistically using Egger's test (the intercept was 0.45 with a 95% CI of -0.48 to 1.38 and a one-tailed p value of 0.16) and the Begg and Mazumdar rank correlation test (Kendall's s value was -0.08 with a one-tailed p value of 0.33). The trim-and-fill method using MIX 2.0 software and comprehensive meta-analysis software did not suggest inputting studies.

3.7 Additional Analysis

Exclusion sensitivity analysis (the effect of exclusion of each study on the summary effect) was performed. The greatest negative shift was after exclusion of the study by Hoffman et al. [45] (mean effect -0.19, 95% CI -0.22 to -0.15), and the greatest positive shift was after exclusion of the study by Valtonen et al. [55] (mean effect -0.13, 95% CI -0.17 to -0.08) [see Electronic Supplementary Material Table S2.1 for details]. The subgroup analysis based on sex showed a statistically significant difference between subgroups (Q = 6.57; df = 1; p = 0.01). The mean effect was -0.22 (95% CI -0.26 to -0.18) for male participants, compared with -0.12 (95% CI -0.19 to -0.05) for female participants (see Fig. 2.3). The subgroup analysis based on participants suffering from MDD did not show a statistically significant difference between subgroups (Q = 0.63; df = 1; p = 0.43). The subgroup analysis based on RER ≥1.1 usage as a criterion for a subject’s sufficient effort during a CPET did not show a statistically significant difference between subgroups (Q = 0.002; df = 1; p = 0.96). The slope of the meta-regression of participants’ age on the effect sizes was not statistically significant (the slope was 0.0026 with a p value of 0.18). The slope of the meta-regression of publication year on the effect sizes was not statistically significant (the slope was -0.007 with a p value of 0.24).
**Figure 2.3**

<table>
<thead>
<tr>
<th>References</th>
<th>Sample size</th>
<th>CC [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Female participants</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hollenberg et al. 2003 [46]</td>
<td>634</td>
<td>-0.11 (-0.19; -0.03)</td>
</tr>
<tr>
<td>Chu et al. 2009 [42]</td>
<td>39</td>
<td>-0.18 (-0.47; 0.14)</td>
</tr>
<tr>
<td>Donath et al. 2010 [44]</td>
<td>30</td>
<td>-0.36 (-0.64; 0)</td>
</tr>
<tr>
<td>Raso et al. 2012 [52]</td>
<td>37</td>
<td>-0.05 (-0.37; 0.28)</td>
</tr>
<tr>
<td>Hunter et al. 2013 [47]</td>
<td>25</td>
<td>-0.05 (-0.44; 0.35)</td>
</tr>
<tr>
<td><strong>Synthesis</strong></td>
<td>765</td>
<td>-0.12 (-0.19; -0.05)</td>
</tr>
<tr>
<td><strong>Male participants</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stewart et al. 2003 [54]</td>
<td>53</td>
<td>-0.25 (-0.49; 0.02)</td>
</tr>
<tr>
<td>Kiive et al. 2004 [48]</td>
<td>46</td>
<td>-0.29 (-0.54; 0)</td>
</tr>
<tr>
<td>Valtonen et al. 2009 [55]</td>
<td>2428</td>
<td>-0.22 (-0.26; -0.18)</td>
</tr>
<tr>
<td><strong>Synthesis</strong></td>
<td>2527</td>
<td>-0.22 (-0.26; -0.18)</td>
</tr>
</tbody>
</table>

Forest plot of effect sizes of correlation between depression symptom severity and cardiorespiratory fitness by sex. CC correlation coefficient, CI confidence interval.
4 Discussion

4.1 Summary of the Evidence

The results of this meta-analysis showed a modest correlation \( r = -0.16 \), 95 % CI -0.21 to -0.10) between depression symptom severity and CRF in healthy and depressed individuals. This correlation differed significantly between sexes \( p = 0.01 \), appearing stronger in male participants \( r = -0.22 \), 95 % CI -0.26 to -0.18) and weaker in female participants \( r = -0.12 \), 95 % CI -0.19 to -0.05). There were no significant findings in the other subgroup analyses and the meta-regression analyses, although it should be mentioned that these analyses had low statistical power due to the small number of studies within subgroups.

4.2 Relevance of the Evidence

The evidence from the included studies is quite relevant to the review question, which is whether a correlation between depression symptom severity and CRF exists, because this was the exact piece of information that was requested from the studies’ authors. The included studies were sufficient to produce a statistically significant summary effect, and they comprised both sexes, a wide spectrum of age and a broad range of depression symptom severity, since both healthy and depressed participants were included. Therefore, the applicability of the findings should be acceptable.

4.3 Quality of the Evidence

Sixteen studies (totaling 4039 participants) of various types were included in the meta-analysis. In three of 11 applicable studies, the participants or the data collectors were blinded to allocation, while in five studies, RER \( \geq 1.1 \) was used as a termination criterion for the CPET. It should be noted that this criterion should be applied routinely [34] but particularly in patients with MDD, in whom mood and somatic symptoms may affect the perception of fatigue [57]. However, it should also be noted that studies assessing CRF without using a symptom-limited CPET were excluded from the meta-analysis; therefore, the risk of measurement error had
already been significantly reduced. Treadmill and cycle ergometer were used in the CPET in half of the studies, respectively. The $\text{VO}_{2\text{peak}}$ derived by a treadmill test is 10% - 20% higher than the $\text{VO}_{2\text{peak}}$ derived by a cycle ergometer test [58] and this might influence the correlation coefficient between CRF and depression. Six studies used researcher-completed depression scales; it should be noted that self-administered scales could be a source of bias. Seven different depression scales were included in the analysis. Although all these scales are validated, they have different ranges of scores, which might also influence the correlation coefficient between the scale score and CRF. Apart from the difference in the effect sizes between studies with male and female participants, the effect sizes were consistent, while exclusion sensitivity analysis did not show large variations in the summary effect. Taking into consideration these remarks, it could be stated that the quality of the evidence renders the findings of the meta-analysis valid, but the reader should be cautious when interpreting them.

4.4 Bias in the Review Process

It was not possible to assess the effects of potential moderators of the correlation between depression symptom severity and CRF — such as age, sex, body mass index (BMI), PA and medication — at the within-study level, because raw data were not available for every study. Between-studies moderator analysis was possible only for sex and age, using subgroup analysis and meta-regression, respectively.

Across-studies bias should also be taken into consideration. Only English-language studies were included in this analysis. Also, although publication bias was not significant, it should be noted that three published studies [59–61] and two unpublished theses [62, 63] were excluded because it was impossible to acquire the relevant data.
4.5 Implications of the Results

The inverse correlation between depression symptom severity and CRF has clinical and prognostic implications. Nearly one fifth of patients who present to primary health care centres suffer from MDD [64]. It is therefore important for clinicians to consider that individuals with MDD may tend to be unfit and thus tend to have an increased risk of all-cause mortality [65, 66], since this may influence assessment and treatment of a significant proportion of patients. It should be noted, though, that the mortality risk due to poor CRF may be included in the already well-known mortality risk associated with MDD, which has been reported in community populations [65, 66] and clinical populations [8, 9]. Twenty-five percent of the global population are not meeting the minimum recommendations for PA [67], while 17 % are physically inactive [68] and therefore in poor CRF. It is equally important for clinicians to consider that unfit individuals may tend to suffer from depression, since this may influence assessment and treatment of this population, as well. Clinicians might consider, for example, screening for depression or might anticipate challenges in adherence to treatment—a common finding in depression [10].

The correlation between depression symptom severity and CRF—especially in men, in whom it is stronger—provides important information regarding treating depression and improving CRF. Regardless of whether there is unidirectional causation, bidirectional causation or influence of other variable(s), it supports the need for research on the effects of treating one on the other. Improving CRF, for example, may reduce depression symptom severity, while treating depression may improve CRF. Regarding the former, exercise is already considered one of the recommended therapeutic modalities for depression [13–16] and, especially in individuals with pathologically low CRF, such as patients with coronary heart disease and heart failure, exercise seems to improve depression and depression-related increased mortality, and this improvement correlates with the improvement in CRF [69, 70]. However, more research on the frequency, intensity, duration and type of exercise should be conducted in order to maximize its beneficial effects. Regarding the
latter, pathologically low CRF is found in patients with comorbid depression (e.g. patients with heart disease), who were excluded from this meta-analysis, and thus this correlation may or may not apply. Two studies [71, 72] have examined the effect of treating depression after myocardial infarction on cardiac outcomes, such as cardiac death and recurrent myocardial infarction, but CRF was not assessed in either of them. There is a clear need for research on this topic.

The correlation between depression symptom severity and CRF suggests that there is a link between depressive symptomatology and an objectively measured biological marker that is a strong predictor of all-cause mortality [21, 22]. Although it is well established that depression is associated with increased all-cause and cardiac mortality [65], the underlying mechanisms are unclear. Potential hypotheses for this association have been suggested, including a diathesis–stress model, reduced heart rate variability, and serotonin and platelet aggregation [73]. If there is a causal relationship between depression and CRF, these new biological pathways need to be explored. In order for causality to be assessed, potential moderators, such as PA, should be controlled for in future trials.

4.6 Potential Interpretations of the Results

Explaining the stronger correlation between depression symptom severity and CRF in men is not straightforward. It should be taken into account that PA is a strong moderator of CRF in healthy individuals [20]. A plausible explanation for the negative correlation between depression symptom severity and CRF is that depressed individuals tend to be less active than non-depressed individuals, therefore engaging in less PA [74] and having lower CRF. Conversely, individuals who are physically active, and therefore have high CRF, may benefit from the antidepressive effects of exercise [13–16] and exhibit low levels of depressive symptoms. Indeed, there is some evidence of an inverse association between depression symptom severity and PA [75, 76], which may be complex and dynamic across time [77]. If PA in women is irrelevant or more weakly correlated with depression symptom severity than it is in men, this would partially explain the weaker
correlation between depression symptom severity and CRF observed in women in this meta-analysis. Indeed, the correlation between depression symptom severity and PA has been reported to be somewhat stronger in men [77, 78].

Another possible contributor to the correlation between depression symptom severity and CRF could be that depressed individuals may actually feel that they have reached fatigue during stress testing sooner than normal because of somatic symptoms of depression, such as muscle soreness, weakness and low back pain. If the prevalence of somatic symptoms in men with depressive symptoms is higher than that in women, this would partially explain the stronger correlation between depression symptom severity and CRF in men in this meta-analysis. However, women have been reported to exhibit more and stronger somatic symptoms than men, in both community populations [79, 80] and clinical populations [81, 82], although this finding is inconsistent, since other studies have failed to confirm this difference, both in community settings [83, 84] and clinical settings [85]. Nevertheless, to the best of our knowledge, there are no studies reporting more or stronger somatic symptoms in men than in women; therefore, this factor does not seem to play a significant role in the sex differences noted in the correlation between depression symptom severity and CRF.

A biological explanation for the stronger correlation between depression symptom severity and CRF in men may be more plausible. There is little dispute about the higher prevalence of MDD in women, which is reported to be twice that in men [86, 87]. This female predisposition is yet unexplained, inconsistent and seems to be determined by social, psychological and biological factors acting simultaneously [88, 89]. Biological factors include genetic predisposition, sex differences in neurotransmitter systems and functional brain asymmetry, and hormonal factors [90]. CRF is higher in men, and even relative CRF (taking into account body weight), which was the form of CRF included in the meta-analysis, is reported to be around 20 % higher in men [91, 92]. Assuming that women may exhibit greater depressive symptoms and lower CRF than men, and men the opposite, may provide an explanation for the overall summary effect, whereas the difference in
the effect size between sexes might be attributed to larger variations in depression symptom severity and/or CRF in women. Since it may be assumed that PA is the strongest moderator in the correlation between depression symptom severity and CRF, these potential biological links could be assessed by controlling for PA in prospective studies. More research is warranted in this promising area.

5 Conclusions

Depression symptom severity is inversely correlated with CRF, and this correlation is stronger in men than in women. Apart from its clinical and prognostic implications, the correlation should stimulate further research on the effects of treating one variable on the other. It may also provide an indication of a biological link between depression and CRF, although this needs to be confirmed by controlling for potential moderators, such as PA.

Compliance with Ethical Standards

Funding: No sources of funding were used in the preparation of this article.

Conflict of interest: Theodoros Papasavvas, Robert O. Bonow, Mohammad Alhashemi and Dominic Micklewright declare that they have no conflicts of interest that are relevant to the content of this article.

Electronic supplementary material: The online version of this article (doi:10.1007/s40279-015-0409-5) contains supplementary material, which is available to authorized users.
References


Chapter 3

Is there an association between depression symptom severity and CRF in patients with heart disease?
Association between depressive symptoms and exercise capacity in patients with heart disease – a meta-analysis

Theodoros Papasavvas, MSc; Mohammad Alhashemi, MD; Dominic Micklewright, PhD

Abstract

Purpose: Depression and reduced exercise capacity are risk factors for poor prognosis in patients with heart disease, but the relationship between the 2 is unclear. We assessed the relationship between depressive symptoms and exercise capacity in patients with heart disease.

Methods: PubMed, Cochrane Library, Google Scholar, and ProQuest databases were browsed for English-language studies published from January 2000 to September 2013. Studies including adult patients with coronary artery disease, heart failure, congenital heart disease, and implantable cardioverter defibrillator, reporting correlation between a depression scale and exercise capacity (VO$_{2peak}$, peak watts, estimated metabolic equivalents, and incremental shuttle walk test distance), as well as studies from which such a correlation could be calculated and provided by the authors, were included. Correlation coefficients (CCs) were converted to Fischer z values, and the analysis was performed using a random-effects model. Then, summary effects and 95% CIs were converted back to CCs.

Results: Fifty-nine studies (25,733 participants) were included. Depressive symptoms were inversely correlated to exercise capacity (CC = −0.15; 95% CI, −0.17 to −0.12). Heterogeneity was significant ($I^2 = 64%$; $Q = 20$).

Author Affiliations: Cardiac Rehabilitation Department, Heart Hospital, Hamad Medical Corporation, Doha, Qatar (Mr. Papasavvas and Dr. Alhashemi); and School of Biological Sciences, University of Essex, Colchester, Essex, United Kingdom (Mr. Papasavvas and Dr. Micklewright).

Supplemental digital content is available for this article. Direct URL citation appears in the printed text and is provided in the HTML and PDF versions of this article on the journal’s Web site (www.jcrpjournal.com).

The authors declare no conflicts of interest.
There was no evidence of publication bias (Fail-safe N = 4681; Egger test: \( P = .06 \); Kendall test: \( P = .29 \)).

**Conclusions:** Patients with heart disease and elevated depressive symptoms may tend to have reduced exercise capacity, and vice versa. This finding has clinical and prognostic implications. It also encourages research on the effects of improving depression on exercise capacity, and vice versa. The effects of potential moderators need to be explored.

**INTRODUCTION**

Cardiovascular diseases are the number one cause of death globally.\(^2\) In 2010, approximately 800,000 people died of cardiovascular diseases in the United States, accounting for one-third of all deaths.\(^3\) The major modifiable contributors to cardiovascular disease mortality are high blood pressure, tobacco use, unhealthy diet and obesity, insufficient physical activity, and abnormal glucose levels.\(^4\) Other risk factors associated with heart disease include dyslipidemia and harmful use of alcohol, \(^1\) as well as emerging risk factors such as depression\(^5\) and anxiety.\(^6\)

The abovementioned risk factors also worsen prognosis in patients with established heart disease or after acute coronary syndrome. Depression, which is evident in nearly 20% of these patients,\(^7\) is associated with a 2.25-fold risk of all-cause mortality after myocardial infarction\(^8\) and has been recommended by the American Heart Association as a risk factor for poor prognosis following acute coronary syndrome.\(^9\) The risk for all-cause mortality in patients suffering from heart failure is increased by 51% in the presence of depression.\(^10\)

Insufficient physical activity leads to poor fitness, which, in patients with established heart disease or after acute coronary syndrome, is also associated with poor prognosis. Fitness can be objectively assessed by determining exercise capacity using a symptom-limited exercise test. Low exercise capacity in these patients is associated with a more than 2-fold risk of all-cause mortality compared with high exercise capacity.\(^11\) Given
the abovementioned clinical significance of depression and exercise capacity in this population, expected outcomes for patients enrolled in cardiac rehabilitation programs include the absence of clinically significant depressive symptoms and a clinically significant increase in exercise capacity. A potential interaction between these 2 variables may therefore be of substantial clinical importance in assessment, risk stratification, goal setting, therapeutic intervention, and rehabilitation. Studies assessing a correlation between depressive symptoms and exercise capacity in patients suffering from heart disease are few; however, this correlation can be assessed in every study that reports measures of depressive symptoms and exercise capacity, regardless of whether this is the purpose of the study or not. We therefore conducted a meta-analysis including studies from which correlations between depressive symptoms and exercise capacity could be calculated.

METHODS

ELIGIBILITY CRITERIA

Studies published between January 2000 and September 2013 were inspected. The restriction to recent studies was decided on the rationale that access to the raw database of the studies in order to calculate correlations would be needed and the availability of such databases would diminish with time. Experimental and observational studies on adult patients with coronary artery disease, heart failure, congenital heart disease, and implantable cardioverter defibrillator (ICD), alone and with comorbid depression (mild, moderate, or major), published and unpublished in English were eligible for inclusion. Studies on patients with comorbid depression were eligible in order for the total range of depressive symptoms to be used for calculating the correlation between depressive symptoms and exercise capacity. Studies that reported correlation(s) between depressive symptoms and exercise capacity, as well as studies from which such correlations could be calculated, were included. Depressive symptoms were measured using validated depression scales. Exercise capacity was measured (1) in the case of cardiopulmonary exercise testing, using
a symptom-limited exercise test and reported as peak oxygen uptake (VO$_2$peak); (2) with treadmill exercise testing, estimated metabolic equivalents (METs) at maximum workload or in watts with cycle ergometer testing; and (3) total distance covered in the case of incremental shuttle walk testing (ISWT). Heart transplantation studies were excluded.

**INFORMATION SOURCES**

A systematic computerized literature search in the databases of PubMed, Cochrane Library, Google Scholar, and Pro-Quest, covering articles from January 2000 to September 2013, was conducted. ProQuest contains unpublished theses, the inclusion of which reduces publication bias. References of included studies were also reviewed, and principal investigators of included studies were asked to provide relevant published or unpublished studies.

**SEARCH**

The titles and abstracts (PubMed), titles (Google Scholar), abstracts (ProQuest), and full text (Cochrane Library) of studies were searched using the following search terms: exercise AND depress*; card* AND depress*; heart AND depress*; stress test* AND depress*; aerobic capacity AND depress*; fitness AND depress*; oxygen uptake AND depress*; VO$_2$$_{\text{max}}$ AND depress*; VO$_2$peak AND depress*; and rehabilitation AND depress*.

All search terms were followed by NOT menopause NOT pregnant NOT postpartum NOT postnatal NOT fibromyalgia NOT Parkinson’s NOT stroke NOT sclerosis NOT spinal cord NOT lumbar NOT low back NOT spondylitis NOT osteoarthritis NOT fracture NOT musculoskeletal NOT cancer NOT bipolar NOT binge NOT adolescent* NOT rat, except in the Google Scholar search.
STUDY SELECTION AND DATA COLLECTION

All potentially relevant studies, judged from title and abstract, were either acquired in full text or their authors were contacted and asked whether an exercise capacity metric (VO_{2peak}, METs, maximum watts, and total distance covered in the ISWT) and a depression scale score had been reported. Studies reporting one of the abovementioned exercise capacity metrics and a depression scale score or correlation between them were included. Authors of such studies were contacted, and the raw data or the correlations were requested.

DATA ITEMS

The correlation between depressive symptoms and exercise capacity was the only data item and was either reported or calculated from the included studies. One correlation coefficient per study was calculated and included in the analysis. If a study had more than 1 group, the correlation for the total cohort was calculated. If a study had more than 1 measurement, the correlation at baseline was calculated.

RISK OF BIAS IN INDIVIDUAL STUDIES

To reduce the risk of selective outcome reporting and to eliminate the bias of lack of blinding of the participants and researchers in prospective studies, the correlation between depressive symptoms and exercise capacity at baseline only was reported or calculated.

SUMMARY MEASURES

Pearson (or Spearman, in case of nonparametric data) correlation coefficients between a depression scale score and an exercise capacity metric were used to create the summary measure. A random-effects model was used, since there is no current knowledge of the effect of moderators (e.g., gender, age) on the true effect size.
PLANNED METHOD OF ANALYSIS

The method of Hedges and Vevea\textsuperscript{13} was used for synthesizing the correlation coefficients. In this method, the correlations are converted to the Fisher z scale and the analysis is performed using the transformed values. Then, the results, such as summary effect and 95% CIs, are converted back to correlations. This method of converting correlations to Fisher z values was chosen to eliminate the bias in the correlation variance, since it depends strongly on the correlation itself.

Any excessive inconsistency (between-studies variation) is considered to reflect true differences between studies and is termed heterogeneity. Heterogeneity was assessed using the Cochran Q $\chi^2$ test and using $I^2$, which represents the percentage of the total variation that is attributed to true differences between studies. Confidence (uncertainty) intervals for $I^2$ were also calculated. MIX 2.0 (Biostat, Englewood, NJ) and Comprehensive Meta-Analysis software (Biostat, Englewood, NJ) were used for statistical analysis.

RISK OF BIAS ACROSS STUDIES

Publication bias was assessed visually on a funnel plot (effect size by the inverse of its standard error) and statistically using the Egger test, the Begg and Mazumdar rank correlation test, and the Trim and Fill method. It should be mentioned, however, that other factors such as study quality and true heterogeneity can produce asymmetry in funnel plots. Duplication was avoided by juxtaposing author names, study sample size, gender, and age. In cases of ambiguity, the authors were contacted. Availability bias (selective inclusion of studies that are easily accessible to the researcher), with regard to omitting studies from which the required data were difficult to acquire, was addressed by contacting all authors of the study, by regular (biweekly) follow-ups, and by collecting data for a period of 12 months. The Rosenthal Fail-safe N statistic was also used to calculate the number of missing studies that would render the summary effect nonsignificant.
ADDITIONAL ANALYSES

Prespecified exclusion sensitivity analysis was performed to assess the effect of exclusion of each study (one at a time) on the summary effect. Post hoc subgroup analyses were performed using a Q-test based on analysis of variance for random-effects model with pooled estimate of \( \tau^2 \) (between-studies dispersion) with the following moderators: gender, with 2 levels: male and female; disease with 2 levels: coronary artery disease and heart failure; exercise test with 4 levels: cardiopulmonary exercise test, simple treadmill test, watts-based cycle ergometer test, and ISWT. Prespecified meta-regression was performed to assess the effect of age on the effect sizes.

RESULTS

STUDY SELECTION

A total of 59 studies were included in the analysis. The search of all databases provided a total of 23,508 citations. Of these, 21,627 studies were discarded after reading the title and/or abstract, since it was clear that they did not meet the criteria. The remaining 1881 studies were acquired in full text or their authors were contacted and asked whether depression and exercise capacity had been assessed. Of these, 1543 and 229 studies were discarded because there was no assessment of depression and/or exercise capacity and because exercise capacity was measured using a submaximal test (e.g. 6-minute walk test), respectively. Of the remaining 109 eligible studies, 16 studies were duplicates or shared the same sample whereas 34 studies were not included because it was impossible to acquire the data from their authors (Figure 3.1).

STUDY CHARACTERISTICS

Fifty-eight published studies and 1 unpublished thesis\textsuperscript{14} were included, totaling 25,733 participants. Eleven studies reported a correlation between depressive symptoms and exercise capacity, and authors of the remaining 48 studies were contacted and provided the raw data or the correlations. Studies were
Flow diagram of study selection process

**Search terms results**
Pubmed ($n = 15382$), Google Scholar ($n = 4343$), ProQuest ($n = 3700$), Cochrane Library ($n = 83$)
Total ($n = 23508$)

Reading title and abstract

Irrelevant studies ($n = 21627$)

Relevant studies ($n = 1881$)

Acquiring full text and/or contacting authors

Studies without depression and/or exercise capacity assessment ($n = 1543$)

Studies without symptom-limited exercise testing ($n = 229$)

Eligible studies ($n = 109$)

Duplicate studies and studies sharing sample ($n = 16$)

Included studies ($n = 59$)

Studies from which data was impossible to acquire ($n = 34$)
experimental and observational, and the majority included patients with coronary artery disease and heart failure. One study included patients with congenital heart disease,\textsuperscript{15} and 2 studies included patients with ICD\textsuperscript{16,17} (Table 3.1).
<table>
<thead>
<tr>
<th>Author</th>
<th>n</th>
<th>Age, mean, y</th>
<th>Gender</th>
<th>Study type</th>
<th>Disease</th>
<th>Exercise test</th>
<th>Scale</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ades et al. (2002)</td>
<td>51</td>
<td>70.6</td>
<td>Both</td>
<td>Cross-sectional</td>
<td>CAD</td>
<td>CPET</td>
<td>GDS</td>
<td>Reported in the paper.</td>
</tr>
<tr>
<td>Ades et al. (2003)</td>
<td>41</td>
<td>72.3</td>
<td>Female</td>
<td>RCT</td>
<td>CAD</td>
<td>CPET</td>
<td>GDS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Asbury et al. (2012)</td>
<td>40</td>
<td>65.1</td>
<td>Both</td>
<td>RCT</td>
<td>CAD</td>
<td>ISWT</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Beckie et al. (2011)</td>
<td>252</td>
<td>63.5</td>
<td>Female</td>
<td>RCT</td>
<td>CAD</td>
<td>TET</td>
<td>CES-D</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Benzer et al. (2007)</td>
<td>241</td>
<td>56</td>
<td>Both</td>
<td>Trial</td>
<td>CAD</td>
<td>CET (watts)</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Bettencourt et al. (2005)</td>
<td>94</td>
<td>57.4</td>
<td>Both</td>
<td>RCT</td>
<td>CAD</td>
<td>TET</td>
<td>BDI</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Bunevicius et al. (2011)</td>
<td>1470</td>
<td>57</td>
<td>Both</td>
<td>Cross-sectional</td>
<td>CAD</td>
<td>CET (watts)</td>
<td>HADS</td>
<td>Reported in the paper.</td>
</tr>
<tr>
<td>De Schutter et al. (2011)</td>
<td>529</td>
<td>64.1</td>
<td>Both</td>
<td>Longitudinal</td>
<td>CAD</td>
<td>CPET</td>
<td>KSQ-D</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Dougherty et al. (2008)</td>
<td>10</td>
<td>54.8</td>
<td>Both</td>
<td>Trial</td>
<td>ICD</td>
<td>CPET</td>
<td>CES-D</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Duarte Freitas et al. (2011)</td>
<td>101</td>
<td>65</td>
<td>Both</td>
<td>Trial</td>
<td>CAD</td>
<td>CET (watts)</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Evangelista et al. (2008)</td>
<td>150</td>
<td>55</td>
<td>Both</td>
<td>Cross-sectional</td>
<td>HF</td>
<td>CPET</td>
<td>BDI</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Faulkner et al. (2013)</td>
<td>22</td>
<td>62</td>
<td>Both</td>
<td>RCT</td>
<td>CAD</td>
<td>ISWT</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Freyssin et al. (2012)</td>
<td>24</td>
<td>54</td>
<td>Both</td>
<td>RCT</td>
<td>HF</td>
<td>CPET</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Gotlieb et al. (2009)</td>
<td>2268</td>
<td>59.3</td>
<td>Both</td>
<td>Cross-sectional</td>
<td>HF</td>
<td>CPET</td>
<td>BDI</td>
<td>Reported in the paper.</td>
</tr>
<tr>
<td>Guck et al. (2012)</td>
<td>36</td>
<td>70.2</td>
<td>Both</td>
<td>Trial</td>
<td>CAD</td>
<td>TET</td>
<td>BDI</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Houchen et al. (2012)</td>
<td>17</td>
<td>67.3</td>
<td>Both</td>
<td>Trial</td>
<td>HF</td>
<td>ISWT</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Hughes et al. (2006)</td>
<td>260</td>
<td>64</td>
<td>Both</td>
<td>Cross-sectional</td>
<td>CAD</td>
<td>TET</td>
<td>BDI</td>
<td>Reported in the paper.</td>
</tr>
<tr>
<td>Hughes et al. (2007)</td>
<td>68</td>
<td>60</td>
<td>Both</td>
<td>RCT</td>
<td>CAD</td>
<td>CPET</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Izawa et al. (2008)</td>
<td>442</td>
<td>61.1</td>
<td>Both</td>
<td>Cross-sectional</td>
<td>CAD</td>
<td>CPET</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Author</td>
<td>n</td>
<td>Age, mean, y</td>
<td>Gender</td>
<td>Study type</td>
<td>Disease</td>
<td>Exercise test</td>
<td>Scale</td>
<td>Correlation</td>
</tr>
<tr>
<td>------------------------------</td>
<td>-----</td>
<td>--------------</td>
<td>--------</td>
<td>------------</td>
<td>---------</td>
<td>---------------</td>
<td>-------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>Jackson et al. (2013)</td>
<td>56</td>
<td>61.1</td>
<td>Both</td>
<td>Trial</td>
<td>CAD</td>
<td>TET</td>
<td>BDI</td>
<td>Reported in the paper.</td>
</tr>
<tr>
<td>Jolly et al. (2007)</td>
<td>359</td>
<td>61.3</td>
<td>Both</td>
<td>RCT</td>
<td>CAD</td>
<td>ISWT</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Jolly et al. (2009)</td>
<td>169</td>
<td>68</td>
<td>Both</td>
<td>RCT</td>
<td>HF</td>
<td>ISWT</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Karapolat et al. (2009)</td>
<td>61</td>
<td>44.6</td>
<td>Both</td>
<td>RCT</td>
<td>HF</td>
<td>CPET</td>
<td>BDI</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Karlsson et al. (2007)</td>
<td>185</td>
<td>63.5</td>
<td>Both</td>
<td>RCT</td>
<td>CAD</td>
<td>CET (watts)</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Kitzman et al. (2010)</td>
<td>53</td>
<td>70</td>
<td>Both</td>
<td>RCT</td>
<td>HF</td>
<td>CPET</td>
<td>CES-D</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Knackstedt et al. (2014)</td>
<td>27</td>
<td>63.7</td>
<td>Both</td>
<td>Trial</td>
<td>HF</td>
<td>CPET</td>
<td>BDI</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Koukouvou et al. (2004)</td>
<td>26</td>
<td>52.5</td>
<td>Both</td>
<td>RCT</td>
<td>HF</td>
<td>CPET</td>
<td>BDI</td>
<td>Not reported in the paper. BDI was preferred to HADS.</td>
</tr>
<tr>
<td>Kulcu et al. (2007)</td>
<td>31</td>
<td>59.3</td>
<td>Both</td>
<td>RCT</td>
<td>HF</td>
<td>CPET</td>
<td>BDI</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Laederach-Hofmann et al. (2007)</td>
<td>25</td>
<td>-</td>
<td>Both</td>
<td>Trial</td>
<td>HF</td>
<td>CPET</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Lavoie et al. (2004)</td>
<td>1367</td>
<td>56.3</td>
<td>Both</td>
<td>Trial</td>
<td>CAD</td>
<td>TET</td>
<td>BDI</td>
<td>Reported in the paper.</td>
</tr>
<tr>
<td>Marchionni et al. (2000)</td>
<td>265</td>
<td>68</td>
<td>Both</td>
<td>Cross-sectional</td>
<td>CAD</td>
<td>CET (watts)</td>
<td>CES-D</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Meyer et al. (2010)</td>
<td>2236</td>
<td>58.6</td>
<td>Both</td>
<td>Longitudinal</td>
<td>CAD</td>
<td>CET (watts)</td>
<td>HADS</td>
<td>Only patients with CAD were included.</td>
</tr>
<tr>
<td>Mourot et al. (2010)</td>
<td>788</td>
<td>56.9</td>
<td>Both</td>
<td>Trial</td>
<td>CAD</td>
<td>CPET</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Mroszczyn-McDonald et al. (2007)</td>
<td>1428</td>
<td>62.3</td>
<td>Both</td>
<td>Trial</td>
<td>CAD</td>
<td>CPET</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Author</td>
<td>n</td>
<td>Age, mean, y</td>
<td>Gender</td>
<td>Study type</td>
<td>Disease</td>
<td>Exercise test</td>
<td>Scale</td>
<td>Correlation</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----</td>
<td>--------------</td>
<td>--------</td>
<td>---------------------</td>
<td>---------</td>
<td>---------------</td>
<td>-------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Muller et al. (2012)</td>
<td>767</td>
<td>25.6</td>
<td>Both</td>
<td>Cross-sectional</td>
<td>CONG</td>
<td>CPET</td>
<td>CES-D</td>
<td>Reported in the paper.</td>
</tr>
<tr>
<td>Reid et al. (2005)</td>
<td>391</td>
<td>58</td>
<td>Both</td>
<td>Cross-sectional</td>
<td>CAD</td>
<td>CPET</td>
<td>CES-D</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Reschke et al. (2000)</td>
<td>99</td>
<td>52.3</td>
<td>Both</td>
<td>Cross-sectional (unpublished)</td>
<td>HF</td>
<td>CPET</td>
<td>BDI</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Rothenbacher et al. (2007)</td>
<td>1176</td>
<td>59</td>
<td>Both</td>
<td>Longitudinal</td>
<td>CAD</td>
<td>CET (watts)</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Ruo et al. (2004)</td>
<td>944</td>
<td>66.8</td>
<td>Both</td>
<td>Cross-sectional</td>
<td>CAD</td>
<td>TET</td>
<td>PHQ9</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Salzwedel et al. (2012)</td>
<td>1519</td>
<td>71.8</td>
<td>Both</td>
<td>Trial</td>
<td>CAD</td>
<td>CET (watts)</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Sanjuan et al. (2013)</td>
<td>91</td>
<td>53.9</td>
<td>Both</td>
<td>Trial</td>
<td>CAD</td>
<td>TET</td>
<td>SCL-90-R-D</td>
<td>Reported in the paper.</td>
</tr>
<tr>
<td>Seki et al. (2003)</td>
<td>38</td>
<td>69.8</td>
<td>Male</td>
<td>RCT</td>
<td>CAD</td>
<td>CPET</td>
<td>SDS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Silberman et al. (2010)</td>
<td>3710</td>
<td>57.7</td>
<td>Both</td>
<td>Trial</td>
<td>CAD</td>
<td>TET</td>
<td>CES-D</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Skobel et al. (2005)</td>
<td>30</td>
<td>56.3</td>
<td>Both</td>
<td>Trial</td>
<td>HF</td>
<td>CPET</td>
<td>BDI</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Smart et al. (2013)</td>
<td>30</td>
<td>63.4</td>
<td>Male</td>
<td>Trial</td>
<td>HF</td>
<td>CPET</td>
<td>CDS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Smiatek et al. (2013)</td>
<td>32</td>
<td>62.2</td>
<td>Male</td>
<td>Trial</td>
<td>ICD</td>
<td>CPET</td>
<td>BDI</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Spaderna et al. (2010)</td>
<td>239</td>
<td>53.5</td>
<td>Both</td>
<td>Cross-sectional</td>
<td>HF</td>
<td>CPET</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Stauber et al. (2012)</td>
<td>728</td>
<td>60.6</td>
<td>Both</td>
<td>Cross-sectional</td>
<td>CAD, HF</td>
<td>CPET</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Steinberg et al. (2011)</td>
<td>55</td>
<td>55.3</td>
<td>Both</td>
<td>Cross-sectional</td>
<td>HF</td>
<td>CPET</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Suzuki et al. (2005)</td>
<td>44</td>
<td>58</td>
<td>Both</td>
<td>Trial</td>
<td>CAD</td>
<td>CPET</td>
<td>SDS</td>
<td>Reported in the paper.</td>
</tr>
<tr>
<td>Swardfager et al. (2008)</td>
<td>366</td>
<td>63.6</td>
<td>Both</td>
<td>Cross-sectional</td>
<td>CAD</td>
<td>CPET</td>
<td>CES-D</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Swardfager et al. (2011)</td>
<td>195</td>
<td>64.3</td>
<td>Both</td>
<td>Trial</td>
<td>CAD</td>
<td>CPET</td>
<td>CES-D</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Author</td>
<td>n</td>
<td>Age, mean, y</td>
<td>Gender</td>
<td>Study type</td>
<td>Disease</td>
<td>Exercise test</td>
<td>Scale</td>
<td>Correlation</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----</td>
<td>--------------</td>
<td>--------</td>
<td>------------</td>
<td>---------</td>
<td>---------------</td>
<td>---------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Turner et al. (2002)</td>
<td>1617</td>
<td>60.7</td>
<td>Both</td>
<td>Trial</td>
<td>CAD</td>
<td>CPET</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Van den Berg-Emons et al. (2004)</td>
<td>28</td>
<td>58.6</td>
<td>Both</td>
<td>RCT</td>
<td>HF</td>
<td>CPET</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Vizza et al. (2007)</td>
<td>176</td>
<td>-</td>
<td>Both</td>
<td>Trial</td>
<td>CAD</td>
<td>TET</td>
<td>CES-D</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Von Kanel et al. (2009)</td>
<td>58</td>
<td>58</td>
<td>Both</td>
<td>Cross-sectional</td>
<td>HF</td>
<td>CET (watts)</td>
<td>HADS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Yeh et al. (2011)</td>
<td>100</td>
<td>67.4</td>
<td>Both</td>
<td>RCT</td>
<td>HF</td>
<td>CPET</td>
<td>POMS</td>
<td>Not reported in the paper.</td>
</tr>
<tr>
<td>Yeh et al. (2013)</td>
<td>16</td>
<td>66</td>
<td>Both</td>
<td>RCT</td>
<td>HF</td>
<td>CPET</td>
<td>POMS</td>
<td>Not reported in the paper.</td>
</tr>
</tbody>
</table>

Characteristics of included studies. BDI, Beck Depression Inventory; CAD, coronary artery disease; CDS, Cardiac Depression Scale; CES-D, Center for Epidemiologic Studies Depression Scale; CET, cycle ergometer test; CPET, cardiopulmonary exercise test; GDS, Geriatric Depression Scale; HADS, Hospital Anxiety and Depression Scale; HF, heart failure; ICD, implantable cardioverter defibrillator; ISWT, incremental shuttle walk test; KSQ-D, Kellner’s Symptom Questionnaire–Depression; n/a, not available; PHQ9, Patient Health Questionnaire–9; POMS, Profile of Mood States questionnaire; RCT, randomized controlled trial; SCL-90-R-D, Revised Symptom Checklist-90-R-Depression; SDS, Zung Self-rating Depression Scale; TET, treadmill exercise test.
**RISK OF BIASES WITHIN STUDIES**

Typical sources of bias within studies, such as lack of concealment of randomization and patient blinding, were not assessed because this meta-analysis used only the baseline data for the total cohort. However, data collector blinding at baseline was assessed because it may affect baseline measurements. Baseline measurements that were conducted before randomization were considered collector blind. Of the 26 studies in which collector blinding was applicable, 15 included collector blinding. The results can be seen in Table 3.2.

**RESULTS OF INDIVIDUAL STUDIES**

Effect sizes, including sample size, 95% CI, z-statistic, and study weights for the individual studies are reported in Table 3.3.
<table>
<thead>
<tr>
<th>Studies</th>
<th>Data collectors blinded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ades et al. (2003)</td>
<td>No</td>
</tr>
<tr>
<td>Asbury et al. (2012)</td>
<td>No</td>
</tr>
<tr>
<td>Beckie et al. (2011)</td>
<td>Yes</td>
</tr>
<tr>
<td>Benzer et al. (2007)</td>
<td>No</td>
</tr>
<tr>
<td>Bettencourt et al. (2005)</td>
<td>No</td>
</tr>
<tr>
<td>Faulkner et al. (2013)</td>
<td>Yes</td>
</tr>
<tr>
<td>Freyssin et al. (2012)</td>
<td>No</td>
</tr>
<tr>
<td>Hughes et al. (2007)</td>
<td>Yes</td>
</tr>
<tr>
<td>Jolly et al. (2007)</td>
<td>Yes</td>
</tr>
<tr>
<td>Jolly et al. (2009)</td>
<td>Yes</td>
</tr>
<tr>
<td>Karapolat et al. (2009)</td>
<td>Yes</td>
</tr>
<tr>
<td>Karlsson et al. (2007)</td>
<td>Yes</td>
</tr>
<tr>
<td>Kitzman et al. (2010)</td>
<td>Yes</td>
</tr>
<tr>
<td>Knackstedt et al. (2014)</td>
<td>No</td>
</tr>
<tr>
<td>Koukouvou et al. (2004)</td>
<td>Yes</td>
</tr>
<tr>
<td>Kulcu et al. (2007)</td>
<td>Yes</td>
</tr>
<tr>
<td>Lavoie et al. (2004)</td>
<td>Yes</td>
</tr>
<tr>
<td>Mourot et al. (2010)</td>
<td>Yes</td>
</tr>
<tr>
<td>Reid et al. (2005)</td>
<td>Yes</td>
</tr>
<tr>
<td>Sanjuan et al. (2013)</td>
<td>No</td>
</tr>
<tr>
<td>Seki et al. (2003)</td>
<td>No</td>
</tr>
<tr>
<td>Stauber et al. (2012)</td>
<td>No</td>
</tr>
<tr>
<td>Swardfager et al. (2011)</td>
<td>No</td>
</tr>
<tr>
<td>Van den Berg-Emons et al. (2004)</td>
<td>No</td>
</tr>
<tr>
<td>Yeh et al. (2011)</td>
<td>Yes</td>
</tr>
<tr>
<td>Yeh et al. (2013)</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Risk of bias within studies
<table>
<thead>
<tr>
<th>Sources</th>
<th>n</th>
<th>CC</th>
<th>CI-</th>
<th>CI+</th>
<th>z</th>
<th>P</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ades et al. (2002)</td>
<td>51</td>
<td>0.01</td>
<td>-0.27</td>
<td>0.28</td>
<td>0.07</td>
<td>0.94</td>
<td>0.72%</td>
</tr>
<tr>
<td>Ades et al. (2003)</td>
<td>41</td>
<td>-0.48</td>
<td>-0.69</td>
<td>-0.20</td>
<td>-3.22</td>
<td>&lt;.01</td>
<td>0.59%</td>
</tr>
<tr>
<td>Asbury et al. (2012)</td>
<td>40</td>
<td>-0.08</td>
<td>-0.38</td>
<td>0.24</td>
<td>-0.49</td>
<td>0.63</td>
<td>0.58%</td>
</tr>
<tr>
<td>Beckie et al. (2011)</td>
<td>252</td>
<td>-0.02</td>
<td>-0.14</td>
<td>0.10</td>
<td>-0.32</td>
<td>0.75</td>
<td>2.18%</td>
</tr>
<tr>
<td>Benzer et al. (2007)</td>
<td>241</td>
<td>-0.15</td>
<td>-0.27</td>
<td>-0.02</td>
<td>-2.33</td>
<td>0.02</td>
<td>2.13%</td>
</tr>
<tr>
<td>Bettencourt et al. (2005)</td>
<td>94</td>
<td>-0.35</td>
<td>-0.52</td>
<td>-0.16</td>
<td>-3.49</td>
<td>&lt;.01</td>
<td>1.19%</td>
</tr>
<tr>
<td>Bunevicius et al. (2011)</td>
<td>1470</td>
<td>-0.08</td>
<td>-0.13</td>
<td>-0.03</td>
<td>-3.07</td>
<td>&lt;.01</td>
<td>3.64%</td>
</tr>
<tr>
<td>De Schutter et al. (2011)</td>
<td>529</td>
<td>-0.12</td>
<td>-0.20</td>
<td>-0.04</td>
<td>-2.77</td>
<td>0.01</td>
<td>2.93%</td>
</tr>
<tr>
<td>Dougherty et al. (2008)</td>
<td>10</td>
<td>-0.24</td>
<td>-0.76</td>
<td>0.46</td>
<td>-0.65</td>
<td>0.52</td>
<td>0.12%</td>
</tr>
<tr>
<td>Duarte Freitas et al. (2011)</td>
<td>101</td>
<td>-0.06</td>
<td>-0.25</td>
<td>0.14</td>
<td>-0.59</td>
<td>0.55</td>
<td>1.25%</td>
</tr>
<tr>
<td>Evangelista et al. (2008)</td>
<td>150</td>
<td>-0.23</td>
<td>-0.38</td>
<td>-0.07</td>
<td>-2.84</td>
<td>&lt;.01</td>
<td>1.63%</td>
</tr>
<tr>
<td>Faulkner et al. (2013)</td>
<td>22</td>
<td>-0.29</td>
<td>-0.63</td>
<td>0.15</td>
<td>-1.30</td>
<td>0.18</td>
<td>0.32%</td>
</tr>
<tr>
<td>Freyssin et al. (2012)</td>
<td>24</td>
<td>0.41</td>
<td>0.01</td>
<td>0.70</td>
<td>2.00</td>
<td>0.01</td>
<td>0.35%</td>
</tr>
<tr>
<td>Gotlieb et al. (2009)</td>
<td>2268</td>
<td>-0.06</td>
<td>-0.10</td>
<td>-0.02</td>
<td>-2.86</td>
<td>&lt;.01</td>
<td>3.83%</td>
</tr>
<tr>
<td>Guck et al. (2012)</td>
<td>36</td>
<td>-0.32</td>
<td>-0.59</td>
<td>0.01</td>
<td>-1.91</td>
<td>0.06</td>
<td>0.52%</td>
</tr>
<tr>
<td>Houchen et al. (2012)</td>
<td>17</td>
<td>-0.45</td>
<td>-0.77</td>
<td>0.04</td>
<td>-1.81</td>
<td>0.07</td>
<td>0.24%</td>
</tr>
<tr>
<td>Hughes et al. (2006)</td>
<td>260</td>
<td>-0.18</td>
<td>-0.30</td>
<td>-0.06</td>
<td>-2.92</td>
<td>&lt;.01</td>
<td>2.21%</td>
</tr>
<tr>
<td>Hughes et al. (2007)</td>
<td>68</td>
<td>-0.02</td>
<td>-0.26</td>
<td>0.22</td>
<td>-0.16</td>
<td>0.87</td>
<td>0.92%</td>
</tr>
<tr>
<td>Izawa et al. (2008)</td>
<td>442</td>
<td>-0.01</td>
<td>-0.10</td>
<td>0.08</td>
<td>-0.21</td>
<td>0.83</td>
<td>2.76%</td>
</tr>
<tr>
<td>Jackson et al. (2013)</td>
<td>56</td>
<td>-0.21</td>
<td>-0.45</td>
<td>0.06</td>
<td>-1.55</td>
<td>0.12</td>
<td>0.78%</td>
</tr>
<tr>
<td>Jolly et al. (2007)</td>
<td>359</td>
<td>-0.29</td>
<td>-0.38</td>
<td>-0.19</td>
<td>-5.63</td>
<td>&lt;.01</td>
<td>2.55%</td>
</tr>
<tr>
<td>Jolly et al. (2009)</td>
<td>169</td>
<td>-0.26</td>
<td>-0.40</td>
<td>-0.11</td>
<td>-3.43</td>
<td>&lt;.01</td>
<td>1.76%</td>
</tr>
<tr>
<td>Karapolat et al. (2009)</td>
<td>61</td>
<td>-0.41</td>
<td>-0.60</td>
<td>-0.18</td>
<td>-3.32</td>
<td>&lt;.01</td>
<td>0.84%</td>
</tr>
<tr>
<td>Karlsson et al. (2007)</td>
<td>185</td>
<td>0.02</td>
<td>-0.12</td>
<td>0.16</td>
<td>0.27</td>
<td>0.79</td>
<td>1.85%</td>
</tr>
<tr>
<td>Kitzman et al. (2010)</td>
<td>53</td>
<td>-0.32</td>
<td>-0.54</td>
<td>-0.05</td>
<td>-2.35</td>
<td>0.02</td>
<td>0.75%</td>
</tr>
<tr>
<td>Knackstedt et al. (2014)</td>
<td>27</td>
<td>-0.05</td>
<td>-0.42</td>
<td>0.34</td>
<td>-0.25</td>
<td>0.81</td>
<td>0.39%</td>
</tr>
<tr>
<td>Koukouvou et al. (2004)</td>
<td>26</td>
<td>-0.30</td>
<td>-0.62</td>
<td>0.10</td>
<td>-1.48</td>
<td>0.14</td>
<td>0.38%</td>
</tr>
<tr>
<td>Kulcu et al. (2007)</td>
<td>31</td>
<td>-0.40</td>
<td>-0.66</td>
<td>-0.05</td>
<td>-2.24</td>
<td>0.02</td>
<td>0.45%</td>
</tr>
<tr>
<td>Laederach-Hofmann et al. (2007)</td>
<td>25</td>
<td>-0.21</td>
<td>-0.56</td>
<td>0.20</td>
<td>-1.00</td>
<td>0.32</td>
<td>0.36%</td>
</tr>
<tr>
<td>Lavoie et al. (2004)</td>
<td>1367</td>
<td>-0.16</td>
<td>-0.21</td>
<td>-0.11</td>
<td>-5.96</td>
<td>&lt;.01</td>
<td>3.61%</td>
</tr>
<tr>
<td>Marchionni et al. (2000)</td>
<td>265</td>
<td>-0.30</td>
<td>-0.41</td>
<td>-0.19</td>
<td>-5.01</td>
<td>&lt;.01</td>
<td>2.23%</td>
</tr>
<tr>
<td>Meyer et al. (2010)</td>
<td>2236</td>
<td>-0.09</td>
<td>-0.13</td>
<td>-0.05</td>
<td>-4.26</td>
<td>&lt;.01</td>
<td>3.82%</td>
</tr>
<tr>
<td>Mourt et al. (2010)</td>
<td>788</td>
<td>-0.18</td>
<td>-0.25</td>
<td>-0.11</td>
<td>-5.10</td>
<td>&lt;.01</td>
<td>3.26%</td>
</tr>
<tr>
<td>Mroszczyk-McDonald et al. (2007)</td>
<td>1428</td>
<td>-0.27</td>
<td>-0.32</td>
<td>-0.22</td>
<td>-10.45</td>
<td>&lt;.01</td>
<td>3.63%</td>
</tr>
<tr>
<td>Muller et al. (2012)</td>
<td>767</td>
<td>-0.16</td>
<td>-0.23</td>
<td>-0.09</td>
<td>-4.46</td>
<td>&lt;.01</td>
<td>3.23%</td>
</tr>
<tr>
<td>Reid et al. (2005)</td>
<td>391</td>
<td>-0.20</td>
<td>-0.29</td>
<td>-0.10</td>
<td>-3.99</td>
<td>&lt;.01</td>
<td>2.64%</td>
</tr>
<tr>
<td>Reschke et al. (2000)</td>
<td>99</td>
<td>-0.22</td>
<td>-0.40</td>
<td>-0.02</td>
<td>-2.19</td>
<td>0.03</td>
<td>1.23%</td>
</tr>
<tr>
<td>Rocha et al. (2011)</td>
<td>92</td>
<td>-0.21</td>
<td>-0.40</td>
<td>-0.01</td>
<td>-2.01</td>
<td>0.04</td>
<td>1.17%</td>
</tr>
<tr>
<td>Rothenbacher et al. (2007)</td>
<td>1176</td>
<td>-0.10</td>
<td>-0.16</td>
<td>-0.04</td>
<td>-3.44</td>
<td>&lt;.01</td>
<td>3.52%</td>
</tr>
<tr>
<td>Ruo et al. (2004)</td>
<td>944</td>
<td>-0.12</td>
<td>-0.18</td>
<td>-0.06</td>
<td>-3.70</td>
<td>&lt;.01</td>
<td>3.38%</td>
</tr>
<tr>
<td>Sources</td>
<td>n</td>
<td>CC</td>
<td>Cl-</td>
<td>Cl+</td>
<td>z</td>
<td>P</td>
<td>Weight</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>-------</td>
<td>--------</td>
</tr>
<tr>
<td>Salzwedel et al. (2012)²²</td>
<td>1519</td>
<td>-0.19</td>
<td>-0.24</td>
<td>-0.14</td>
<td>-7.49</td>
<td>&lt;.01</td>
<td>3.66%</td>
</tr>
<tr>
<td>Sanjuan et al. (2013)²³</td>
<td>91</td>
<td>-0.30</td>
<td>-0.48</td>
<td>-0.10</td>
<td>-2.90</td>
<td>&lt;.01</td>
<td>1.16%</td>
</tr>
<tr>
<td>Seki et al. (2003)²⁴</td>
<td>38</td>
<td>-0.09</td>
<td>-0.40</td>
<td>0.24</td>
<td>-0.53</td>
<td>0.86</td>
<td>0.55%</td>
</tr>
<tr>
<td>Silberman et al. (2010)²⁵</td>
<td>3710</td>
<td>-0.09</td>
<td>-0.12</td>
<td>-0.06</td>
<td>-5.49</td>
<td>&lt;.01</td>
<td>3.97%</td>
</tr>
<tr>
<td>Skobel et al. (2005)²⁶</td>
<td>30</td>
<td>-0.12</td>
<td>-0.46</td>
<td>0.25</td>
<td>-0.63</td>
<td>0.53</td>
<td>0.44%</td>
</tr>
<tr>
<td>Smart et al. (2013)²⁷</td>
<td>30</td>
<td>-0.06</td>
<td>-0.41</td>
<td>0.31</td>
<td>-0.31</td>
<td>0.75</td>
<td>0.44%</td>
</tr>
<tr>
<td>Smiatek et al. (2013)¹⁷</td>
<td>32</td>
<td>-0.15</td>
<td>-0.47</td>
<td>0.21</td>
<td>-0.81</td>
<td>0.42</td>
<td>0.47%</td>
</tr>
<tr>
<td>Spaderna et al. (2010)²⁸</td>
<td>239</td>
<td>-0.04</td>
<td>-0.17</td>
<td>0.09</td>
<td>-0.61</td>
<td>0.54</td>
<td>2.12%</td>
</tr>
<tr>
<td>Stauber et al. (2012)²⁹</td>
<td>728</td>
<td>-0.16</td>
<td>-0.23</td>
<td>-0.09</td>
<td>-4.35</td>
<td>&lt;.01</td>
<td>3.19%</td>
</tr>
<tr>
<td>Steinberg et al. (2011)²⁰</td>
<td>55</td>
<td>-0.10</td>
<td>-0.36</td>
<td>0.17</td>
<td>-0.72</td>
<td>0.47</td>
<td>0.77%</td>
</tr>
<tr>
<td>Suzuki et al. (2005)²¹</td>
<td>44</td>
<td>0.09</td>
<td>-0.21</td>
<td>0.38</td>
<td>0.58</td>
<td>0.56</td>
<td>0.63%</td>
</tr>
<tr>
<td>Swardfager et al. (2008)²²</td>
<td>366</td>
<td>-0.19</td>
<td>-0.29</td>
<td>-0.09</td>
<td>-3.66</td>
<td>&lt;.01</td>
<td>2.57%</td>
</tr>
<tr>
<td>Swardfager et al. (2011)²³</td>
<td>195</td>
<td>0.03</td>
<td>-0.11</td>
<td>0.17</td>
<td>0.42</td>
<td>0.68</td>
<td>1.91%</td>
</tr>
<tr>
<td>Turner et al. (2002)²⁴</td>
<td>1617</td>
<td>-0.16</td>
<td>-0.21</td>
<td>-0.11</td>
<td>-6.48</td>
<td>&lt;.01</td>
<td>3.69%</td>
</tr>
<tr>
<td>Van den Berg-Emons et al. (2004)²⁵</td>
<td>28</td>
<td>-0.04</td>
<td>-0.41</td>
<td>0.34</td>
<td>-0.20</td>
<td>0.84</td>
<td>0.41%</td>
</tr>
<tr>
<td>Vizza et al. (2007)²⁶</td>
<td>176</td>
<td>-0.14</td>
<td>-0.28</td>
<td>0.01</td>
<td>-1.85</td>
<td>0.06</td>
<td>1.80%</td>
</tr>
<tr>
<td>Von Kanel et al. (2009)²⁷</td>
<td>58</td>
<td>-0.11</td>
<td>-0.36</td>
<td>0.15</td>
<td>-0.82</td>
<td>0.41</td>
<td>0.81%</td>
</tr>
<tr>
<td>Yeh et al. (2011)²⁸</td>
<td>100</td>
<td>-0.09</td>
<td>-0.28</td>
<td>0.11</td>
<td>-0.89</td>
<td>0.37</td>
<td>1.24%</td>
</tr>
<tr>
<td>Yeh et al. (2013)²⁹</td>
<td>16</td>
<td>-0.56</td>
<td>-0.83</td>
<td>-0.09</td>
<td>-2.28</td>
<td>0.09</td>
<td>0.22%</td>
</tr>
<tr>
<td>Summary</td>
<td>25733</td>
<td>-0.15</td>
<td>-0.17</td>
<td>-0.12</td>
<td>-11.16</td>
<td>&lt;.01</td>
<td>100%</td>
</tr>
</tbody>
</table>

Details of all effects sizes
SYNTHESES OF RESULTS

The summary correlation coefficient was −0.15 with a 95% CI from −0.17 to −0.12 (Figure 3.2). Heterogeneity was statistically significant ($I^2 = 64\%$ with a 95% CI from 52 to 73; $Q = 161; df = 58; P < .001$).

<table>
<thead>
<tr>
<th>Author</th>
<th>CC (CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ades (2002)</td>
<td>0.01 (0.27, 0.28)</td>
<td>0.94</td>
</tr>
<tr>
<td>Ades (2003)</td>
<td>−0.48 (−0.66, −0.2)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Asbury (2012)</td>
<td>−0.08 (−0.38, 0.24)</td>
<td>0.63</td>
</tr>
<tr>
<td>Beckel (2012)</td>
<td>0.02 (−0.16, 0.12)</td>
<td>0.75</td>
</tr>
<tr>
<td>Benzer (2007)</td>
<td>−0.15 (−0.27, 0.02)</td>
<td>0.02</td>
</tr>
<tr>
<td>Betterancourt (2005)</td>
<td>−0.35 (−0.52, 0.16)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Boves (2011)</td>
<td>−0.08 (−0.13, 0.03)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>De Schutter (2011)</td>
<td>−0.12 (−0.2, −0.04)</td>
<td>0.01</td>
</tr>
<tr>
<td>Dougherty (2003)</td>
<td>−0.24 (−0.76, 0.34)</td>
<td>0.52</td>
</tr>
<tr>
<td>Duarte Freitas (2011)</td>
<td>−0.06 (−0.25, 0.14)</td>
<td>0.35</td>
</tr>
<tr>
<td>Evangelista (2008)</td>
<td>−0.23 (−0.38, 0.07)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Feitl (2015)</td>
<td>−0.29 (−0.65, 0.15)</td>
<td>0.19</td>
</tr>
<tr>
<td>Freyman (2012)</td>
<td>0.41 (0.01, 0.7)</td>
<td>0.05</td>
</tr>
<tr>
<td>Gottlieb (2006)</td>
<td>−0.06 (−0.1, −0.02)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Guecz (2012)</td>
<td>−0.32 (−0.59, 0.01)</td>
<td>0.06</td>
</tr>
<tr>
<td>Houchen (2012)</td>
<td>−0.45 (−0.77, 0.04)</td>
<td>0.07</td>
</tr>
<tr>
<td>Hughes (2009)</td>
<td>−0.18 (−0.4, 0.06)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Hughes (2007)</td>
<td>−0.02 (−0.26, 0.22)</td>
<td>0.87</td>
</tr>
<tr>
<td>Iwawa (2008)</td>
<td>−0.01 (−0.01, 0.08)</td>
<td>0.83</td>
</tr>
<tr>
<td>Jackson (2013)</td>
<td>−0.21 (−0.45, 0.06)</td>
<td>0.12</td>
</tr>
<tr>
<td>Jolly (2007)</td>
<td>−0.29 (−0.38, 0.19)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Jolly (2009)</td>
<td>−0.36 (−0.4, −0.11)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Karpinski (2009)</td>
<td>−0.41 (−0.6, −0.18)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Karlsson (2007)</td>
<td>0.02 (0.12, 0.16)</td>
<td>0.79</td>
</tr>
<tr>
<td>Kitzmann (2010)</td>
<td>−0.32 (−0.54, 0.05)</td>
<td>0.02</td>
</tr>
<tr>
<td>Knackstedt (2014)</td>
<td>−0.05 (−0.42, 0.34)</td>
<td>0.81</td>
</tr>
<tr>
<td>Koulouvaris (2004)</td>
<td>−0.3 (−0.62, 0.1)</td>
<td>0.14</td>
</tr>
<tr>
<td>Kuleu (2007)</td>
<td>−0.4 (−0.66, −0.05)</td>
<td>0.02</td>
</tr>
<tr>
<td>Lauchrich-Hofmann (2007)</td>
<td>−0.21 (−0.56, 0.2)</td>
<td>0.32</td>
</tr>
<tr>
<td>Lavoie (2004)</td>
<td>−0.16 (−0.21, −0.11)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Marchionni (2000)</td>
<td>−0.3 (−0.41, −0.19)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Meyer (2010)</td>
<td>−0.09 (−0.13, 0.05)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Moiriot (2010)</td>
<td>−0.18 (−0.25, 0.11)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Mrozyczki-McDonald (2007)</td>
<td>−0.27 (−0.32, 0.22)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Muller (2012)</td>
<td>−0.16 (−0.28, 0.09)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Reiss (2005)</td>
<td>−0.2 (−0.29, −0.1)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Reschke (2009)</td>
<td>−0.22 (−0.4, −0.02)</td>
<td>0.03</td>
</tr>
<tr>
<td>Rocha (2011)</td>
<td>−0.21 (−0.4, −0.01)</td>
<td>0.04</td>
</tr>
<tr>
<td>Rothhakwacha (2007)</td>
<td>−0.1 (−0.16, −0.04)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Ruo (2004)</td>
<td>−0.12 (−0.18, −0.06)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Salzwedel (2012)</td>
<td>−0.19 (−0.24, −0.14)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Sajnani (2013)</td>
<td>−0.3 (−0.48, −0.1)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Setti (2005)</td>
<td>−0.09 (−0.4, 0.24)</td>
<td>0.59</td>
</tr>
<tr>
<td>Silverman (2010)</td>
<td>−0.09 (−0.12, 0.06)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Stoeber (2005)</td>
<td>−0.12 (−0.46, 0.25)</td>
<td>0.53</td>
</tr>
<tr>
<td>Stroht (2012)</td>
<td>−0.2 (−0.43, 0.31)</td>
<td>0.75</td>
</tr>
<tr>
<td>Snider (2013)</td>
<td>−0.15 (−0.47, 0.21)</td>
<td>0.42</td>
</tr>
<tr>
<td>Spaderna (2010)</td>
<td>−0.04 (−0.17, 0.09)</td>
<td>0.84</td>
</tr>
<tr>
<td>Slabber (2012)</td>
<td>−0.16 (−0.28, −0.09)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Steinberg (2011)</td>
<td>−0.1 (−0.36, 0.17)</td>
<td>0.47</td>
</tr>
<tr>
<td>Streat (2005)</td>
<td>0.09 (−0.21, 0.38)</td>
<td>0.46</td>
</tr>
<tr>
<td>Steward (2008)</td>
<td>−0.19 (−0.26, −0.09)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Steward (2011)</td>
<td>0.03 (−0.01, 0.17)</td>
<td>0.08</td>
</tr>
<tr>
<td>Tarner (2002)</td>
<td>−0.16 (−0.21, −0.11)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Van den Bergh-Ensoms (2004)</td>
<td>−0.04 (−0.41, 0.34)</td>
<td>0.84</td>
</tr>
<tr>
<td>Yalva (2007)</td>
<td>−0.14 (−0.28, 0.03)</td>
<td>0.06</td>
</tr>
<tr>
<td>Von Kaner (2009)</td>
<td>−0.11 (−0.36, 0.15)</td>
<td>0.41</td>
</tr>
<tr>
<td>Yeh (2011)</td>
<td>−0.09 (−0.28, 0.11)</td>
<td>0.37</td>
</tr>
<tr>
<td>Yeh (2013)</td>
<td>−0.46 (−0.85, −0.09)</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Forest plot of effect sizes. CC indicates correlation coefficient.
The visual impression of the heterogeneity funnel plot (effect size by the inverse of its standard error) showed no sign of publication bias (Figure 3.3). This was confirmed statistically using the Egger test (intercept was $-0.51$ with a 95% CI from $-1.15$ to $0.13$, and a 1-tailed $P = .06$) and the Begg and Mazumdar rank correlation test (Kendall $\tau$ was $-0.05$ with a 1-tailed $P = .29$). The Rosenthal Fail-safe N was 4681. The Trim and Fill method using the MIX software did not suggest inputting studies; however, the Comprehensive Meta-Analysis software suggested inputting 4 studies, adjusting the summary effect to $-0.14$ with a 95% CI from $-0.17$ to $-0.12$.

Figure 3.3

Funnel plot of effect size by precision. CC indicates correlation coefficient.
**ADDITIONAL ANALYSES**

Exclusion sensitivity analysis (the effect of exclusion of each study to the summary effect) did not show clinically significant findings. The greatest positive shift was after the exclusion of the studies of Jolly et al\textsuperscript{17} or Mroszczyk-McDonald et al\textsuperscript{18} (summary effect $-0.14$ with a 95% CI from $-0.17$ to $-0.12$), and the greatest negative shift was after the exclusion of several studies and affected only the 95% CI (summary effect $-0.15$ with a 95% CI from $-0.18$ to $-0.13$).

The subgroup analysis based on gender did not show statistically significant results ($Q = 0.05$; df = 1; $P = .8$).

The subgroup analysis based on disease did not show statistically significant results ($Q = 0.07$; df = 1; $P = .8$).

The subgroup analysis based on the exercise test did not show statistically significant results ($Q = 5.3$; df = 3; $P = .15$).

The slope of the meta-regression of participant age on the effect sizes was not statistically significant (the slope was $-0.001$ with a $P$ value of .2).

**DISCUSSION**

The results of this study showed a modest correlation ($r = -0.15$; 95% CI, $-0.17$ to $-0.12$) between depressive symptoms and exercise capacity in patients with heart disease that was irrelevant to gender, disease type, and type of exercise test.

One plausible explanation of this relationship is behavioral. The higher the depressive symptoms, the weaker was the motivation to engage in physical activity, which can result in reduced exercise capacity. Indeed, there is evidence that physical activity is associated with depression\textsuperscript{19} and that depression may be a significant risk factor for developing a sedentary lifestyle.\textsuperscript{20} Only four studies reported objectively measured physical activity levels using pedometer\textsuperscript{73} and accelerometer\textsuperscript{15,53,84}, while 19 studies used various self-reported scales that included physical activity; therefore, it was not possible to assess the effect of physical activity on the
correlation between depressive symptoms and exercise capacity. Another plausible explanation of the relationship between depressive symptoms and exercise capacity is biological. There is strong evidence that exercise improves cognitive abilities and enhances brain neuroplasticity. Levels of brain-derived neurotrophic factor, which is involved in cellular analogues of memory formation, plasticity, and cell proliferation, are reduced in depression and can be increased with aerobic exercise. Moreover, inflammation is present in depression whereas levels of inflammatory markers, such as C-reactive protein, are inversely associated with cardiorespiratory fitness in patients with heart disease.

The association between depressive symptoms and exercise capacity in patients with heart disease has implications regarding patient assessment and treatment plan. Clinicians should consider that patients with poor exercise capacity may tend to have elevated depressive symptoms and thus may consider conducting a more thorough psychosocial evaluation of this patient group; they may also be prepared to anticipate depression-related challenges, such as poor adherence to treatment, including rehabilitation. Furthermore, clinicians may consider elevated depressive symptoms in the differential diagnosis of poor exercise capacity, in addition to other factors, such as age, function of cardiovascular and pulmonary systems, body weight, carbohydrate intake, and physical activity in patients with heart disease. Finally, it may be possible that the reduced exercise capacity observed in patients with elevated depressive symptoms is a result of suboptimal exercise test performance due to poor motivation, which could prevent the detection of potential ischemia. Perhaps, a cardiopulmonary exercise test, in which performance of the subject can be objectively assessed by using the respiratory exchange ratio ≥1.1 as a test termination criterion, would be more appropriate for patients with heart disease and elevated depressive symptoms.

The association between depressive symptoms and exercise capacity in patients with heart disease also has implications regarding patient care. Because of the linear relationship between depressive symptoms and exercise capacity, it should be considered that even elevated depressive symptoms (not major depressive
disorder) may affect exercise capacity. This is of clinical importance because the prevalence of subclinical depression is significantly larger than that of a major depressive disorder in patients with heart disease and it may predict mortality after myocardial infarction. Even small reductions in exercise capacity may affect mood. This is of clinical importance, especially not only for cardiac rehabilitation programs where the majority of patients entering cardiac rehabilitation exhibit poor exercise capacity but also for all patients with heart disease where exercise capacity is a strong predictor of mortality. Conversely, even small increases in exercise capacity may improve mood and reduce depression-related increased mortality in patients with coronary heart disease and heart failure.

Furthermore, the abovementioned association warrants a multidimensional approach to treating depression and improving exercise capacity in patients with heart disease. Regardless of whether there is unidirectional causation, bidirectional causation, or the influence of other variable(s), conducting research on the effects of treating the one on the other is warranted. Improving exercise capacity through exercise interventions may reduce depressive symptoms. There is evidence that exercise is superior to placebo in reducing depressive symptoms in patients with coronary artery disease; however, more research is needed in order for the optimum frequency, intensity, type, and duration of exercise to be determined, as well as for assessing whether the effect is transient or longer-lasting. Reducing depressive symptoms might improve exercise capacity. There is evidence that psychological and pharmacological interventions for depression reduce depressive symptoms in patients with coronary artery disease, but, to our knowledge, their effects on exercise capacity have not been studied.

Finally, the association between depressive symptoms and exercise capacity in patients with heart disease further supports the status of depression as a risk factor for poor prognosis following acute coronary syndrome and it adds to the pool of evidence that links depression to worse prognosis in patients with heart failure, since depression is associated with an established strong and independent predictor of mortality in
this patient group. The abovementioned clinical and prognostic implications of this association also support the American Heart Association’s recommendation for depression screening in patients with coronary heart disease\(^{36}\) and advocates for the extension of this recommendation to patients with heart failure.

**LIMITATIONS**

It was not possible to assess the effects of potential moderators of the association between depressive symptoms and exercise capacity, such as age, gender, body mass index, physical activity, and medication, in within-study level because raw data were not available for every study. Between-study analysis was only possible for gender and age, using subgroup analysis and meta-regression, respectively.

Within- and across-studies bias should be taken into consideration. Blinding of data collectors was performed in 15 of 26 applicable studies. Only studies in English language were included, and, although publication bias was not significant, it should be mentioned that 34 eligible studies (see Supplemental Digital Content 3.1, available at: http://links.lww.com/JCRP/A27) were not included in the analysis because it was impossible to acquire the necessary additional data from their authors.

**CONCLUSIONS**

Patients with heart disease and elevated depressive symptoms may tend to have reduced exercise capacity, and vice versa. This finding has implications for patient assessment, plan of care, treatment, rehabilitation, and prognosis. It also encourages research on the effects of improving depression on exercise capacity, and vice versa. The effects of potential moderators and mediators need to be explored.
REFERENCES

108


Chapter 4

Is the relationship between depression symptom severity and CRF in patients with heart disease moderated?
Moderators of the relationship between depression symptom severity and cardiorespiratory fitness in patients with coronary heart disease

ABSTRACT

OBJECTIVE: To assess the relationship between depression symptom severity and cardiorespiratory fitness (CRF) and its potential moderators in patients with coronary heart disease.

METHODS: Data from the initial assessment of consecutive patients with coronary heart disease (n = 1489) entering a cardiac rehabilitation program from 1 January 1993 to 31 December 2002 were used. The moderating effect of sex, age, cardiovascular diagnosis, presence of diabetes, smoking history, body mass index (BMI) and anxiety on the correlation between depression symptom severity and CRF was assessed using linear and logistic regression analysis.

RESULTS AND CONCLUSIONS: Depression symptom severity was inversely correlated to CRF (r = -0.24, p < 0.001). Sex (p = 0.007) and BMI (p < 0.001) moderated the correlation: Lower BMI enhanced the correlation, while higher BMI attenuated the correlation in male patients and rendered it statistically not significant in female patients. These findings may assist clinicians on patient assessment and clinical management, and may warrant further research on the effects of improving depression symptom severity on CRF and vice versa.
INTRODUCTION

There is substantial evidence that depression is a strong and independent risk factor for the development of heart disease (Frasure-Smith and Lesperance 2005, Nair et al., 2012, Whooley and Wong 2013, Hare et al., 2014, Bradley and Rumsfeld 2015, Carney and Freedland 2017, Turk-Adawi et al., 2018) increasing its risk by 30% - 60% (Van der Kooy et al., 2007, Charlson et al., 2013, Gan et al., 2014). Indeed, approximately 20% of patients with heart disease suffer from major depressive disorder (Rutledge et al., 2006, Shanmugasegaram et al., 2012, Elderon and Whooley 2013, Whooley and Wong 2013). The potential pathways linking depression and heart disease have been the topic of extensive research lately. Several mechanisms have been hypothesized and studied, including serotonergic pathway and platelet dysfunction, inflammation, and autonomic nervous system and hypothalamic-pituitary-adrenal axis imbalance (Sher et al., 2010, Chauvet-Gelinier et al., 2013, Seligman and Nemeroff 2015), as well as stress pathways (Chauvet-Gelinier et al., 2013, Dhar and Barton 2016), changes in arterial structure and its contractile and relaxing functions (Bouzinova et al., 2015), unhealthy lifestyle (Penninx 2017), and genetics (Chauvet-Gelinier, et al. 2013, Seligman and Nemeroff 2015, Penninx 2017). Although the precise pathways linking depression to heart disease are still unclear, the impact of depression on patients with heart disease is clear, increasing mortality by more than 100% (Rutledge et al., 2006, Meijer et al., 2011, Leung et al., 2012, Fan et al., 2014) and establishing depression as an independent risk factor for poor prognosis in patients with acute coronary syndrome (Lichtman et al., 2014).

Low cardiorespiratory fitness (CRF) is a strong and independent risk factor for the development of heart disease (Pandey et al., 2015, Kondamudi et al., 2017, Omar et al., 2018) and respectively, improved CRF is associated with decreased incidence of heart disease (Swift et al., 2013, Lavie et al., 2015, Nayor and Vasan 2015, Al-Mallah et al., 2018). In healthy individuals CRF is strongly associated with all-cause mortality, whereas each 1-MET (MET, approximately 3.5 mL · kg⁻¹ · min⁻¹) increase confers a 12-15% improvement in survival (Myers et al., 2002, Kokkinos et al., 2008, Kodama et al., 2009, Nes et al., 2014, McAuley et al., 2016). In
patients with heart disease high CRF is associated with 55-68% reduction in mortality (Kavanagh et al., 2002, Martin et al., 2013). Consequently, patients with heart disease should be enrolled in exercise-based cardiac rehabilitation (Smith et al., 2011, Piepoli et al., 2016), which increases CRF (Haykowsky et al., 2013, Elliott et al., 2015, Santos et al., 2018) and has been reported to reduce all-cause mortality by 20-64% (Taylor et al., 2004, Colbert et al., 2015, Pouche et al., 2016, Chernomordik et al., 2017, Santiago de Araujo Pio et al., 2017, Bachmann et al., 2018).

The American Heart Association has suggested the routine screening for depression in patients with coronary heart disease (Lichtman et al., 2008) and it also recently suggested the use of CRF as a clinical vital sign (Ross et al., 2016). It is therefore expected that reducing depressive symptoms and improving CRF are important therapeutic goals for this population (Balady et al., 2007, Piepoli et al., 2016). Given the clinical significance of depression and CRF in patients with heart disease, a potential interaction between the two could also be of clinical significance regarding patient assessment and clinical management. Indeed, we found that depressive symptoms and CRF are negatively correlated in both healthy individuals (Papasavvas et al., 2016) and patients with heart disease (Papasavvas et al., 2017); however, in these meta-analyses we were not able to assess the effects of potential moderators of this relationship, such as sex, age, and body mass index (BMI), at the within-study level. Potential moderations of the relationship could better inform clinicians on clinical decision making, while they could also elucidate pathways relating depression to CRF. Additionally, there is a probability that the correlation between depressive symptoms and CRF that we found in our meta-analyses is causative. In order for causality to be tested potential moderators of this correlation should be identified and controlled or taken into consideration in the study protocol. To my knowledge, there is no published study assessing the potential moderators of the relationship between depression symptom severity and CRF, either in healthy individuals or clinical populations. The purpose of this study was to assess the relationship between depression symptom severity and CRF and conduct a moderator analysis in a large sample of patients with coronary heart disease.
METHODS

Setting and participants

This study used the database of an open access study (Barons et al., 2015). Consecutive patients (n = 2714) who were referred for cardiac rehabilitation to the Basingstoke and Alton (Hampshire, UK) cardiac rehabilitation program between 1 January 1993 and 31 December 2002 and were registered with the program were included in that study. Patients were typically enrolled in the program 2-6 weeks after their coronary event (acute myocardial infarction, episode of unstable angina or revascularization) and underwent initial and final assessment, at the beginning and at the end of the program, respectively, which included measurements of physical and psychological indices.

Data

Data from the initial assessment were included in this study. This included sex, age, cardiovascular diagnosis, presence of diabetes, smoking history, BMI, CRF, and depression and anxiety symptom severity. Sex, age, diagnosis, diabetes, smoking history, BMI, and anxiety were assessed as potential moderators of the relationship between depression symptom severity and CRF because they are related with either or both depression and CRF [see indicative references for sex (Laukkanen et al., 2009, Shanmugasegaram et al., 2012), age (Stordal et al., 2003, Laukkanen et al., 2009), diagnosis (Laukkanen et al., 2009), diabetes (Moulton et al., 2015, Zaccardi et al., 2015), smoking (Luger et al., 2014, Siddall et al., 2017), BMI (Laukkanen et al., 2009, Pereira-Miranda et al., 2017), and anxiety (Pollack 2005, Loprinzi et al., 2017)]. CRF was assessed by estimating peak oxygen uptake (VO\text{2PEAK}) either on cycle ergometer or treadmill using either the Bruce protocol (Bruce et al., 1973), or the modified Bruce protocol for frail or elderly patients (Bruce 1971), using the same endpoints. VO\text{2PEAK} from the cycle ergometer was estimated using commonly used formula (Astrand and Rodahl 1986); VO\text{2PEAK} from the Bruce test was estimated assuming that each minute of the Bruce protocol uses 1 MET; VO\text{2PEAK} from the modified Bruce test was estimated assuming that the first three stages of the modified Bruce
protocol each use 1 MET. Depression and anxiety symptom severity were assessed using the Hospital Anxiety and Depression Scale (HADS) (Zigmond and Snaith 1983) depression and anxiety subscale scores, respectively.

Statistical Analysis

After establishing a statistically significant correlation between depression symptom severity and CRF ($r = -0.24, p < 0.001$) moderator analysis was performed using the PROCESS (Hayes 2017) macro for SPSS, a linear and logistic regression path analysis modeling tool. The HADS depression subscale score (HADS-D) was set as predictor variable, $V_2\text{PO}_{\text{peak}}$ was set as outcome variable, and sex, age, cardiovascular diagnosis, presence of diabetes, smoking history, BMI, and anxiety symptom severity were set as moderators. The effect of each moderator was assessed separately, setting all other potential moderators as covariates. Mean centering of the HADS-D and the moderator was performed when zero was not a meaningful value for the moderator. Moderation was established if the interaction between the HADS-D and the moderator was statistically significant and was followed by assessing the conditional effect of the HADS-D on CRF at specific values of the moderator (0, 1, 2, 3 ... etc. for categorical moderator; 16th, 50th, and 84th percentiles for continuous moderator).

RESULTS

Complete data for all variables included in the analysis were available in 1489 patients (see Table 4.1) and respective data for patients entering cardiac rehabilitation in 2006 and 2017 in the UK are also provided from the National Audit of Cardiac Rehabilitation (NACR) 2007 (Lewin 2007) and 2018 (Doherty et al., 2018) reports, respectively. Depression symptom severity was inversely correlated to CRF ($r = -0.24, p < 0.001$). The interaction between depression symptom severity and age ($p = 0.08$), cardiovascular diagnosis ($p = 0.94$), presence of diabetes ($p = 0.83$), smoking history ($p = 0.18$), and anxiety ($p = 0.42$) was not statistically significant. The interaction between depression symptom severity and sex (model A) was statistically significant ($t = -2.37, \beta = -0.3, p = 0.018$) and the interaction between depression symptom severity and BMI
(model B – depression symptom severity, BMI mean-centered) was also statistically significant \( (t = 4.39, \beta = 0.05, p < 0.001) \). Since two moderators were found to be statistically significant, a final model including them both was assessed (model C – depression symptom severity, sex, BMI mean-centered), adjusting \( \alpha \)-level for familywise error rate to \( \alpha = 0.025 \). Both sex \( (t = -3.41, \beta = -0.44, p = 0.007) \) and BMI \( (t = 5.04, \beta = 0.06, p < 0.001) \) were found to be significant moderators of the relationship between depression symptom severity and CRF and this model explained most of the variance \( (R^2 = 0.454, \text{see Table 4.2 for unstandardized coefficients in all models}) \). As it can be seen in Figure 4.1, in male patients lower BMI tends to enhance the negative relationship between depression symptom severity and CRF, while higher BMI tends to attenuate it. In female patients (Figure 4.2), the relationship is statistically significant only in patients with lower BMI.
<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>%</th>
<th>NACR 2007</th>
<th>NACR 2018</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1198</td>
<td>80.5</td>
<td>73</td>
<td>71</td>
</tr>
<tr>
<td>Female</td>
<td>291</td>
<td>19.5</td>
<td>27</td>
<td>29</td>
</tr>
<tr>
<td><strong>Cardiovascular diagnosis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI</td>
<td>831</td>
<td>56</td>
<td>52</td>
<td>16</td>
</tr>
<tr>
<td>Angina</td>
<td>95</td>
<td>6</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Other cardiac conditions</td>
<td>29</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td><strong>Treatment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CABG</td>
<td>318</td>
<td>21</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Elective PCI</td>
<td>146</td>
<td>10</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>Primary PCI</td>
<td>70</td>
<td>5</td>
<td>5</td>
<td>29</td>
</tr>
<tr>
<td><strong>Diabetes</strong></td>
<td>195</td>
<td>13</td>
<td>19</td>
<td>25</td>
</tr>
<tr>
<td><strong>Smoking history</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never smoked</td>
<td>435</td>
<td>29</td>
<td>86% non smokers</td>
<td>94% non smokers</td>
</tr>
<tr>
<td>Not smoked for &gt; 10 years</td>
<td>373</td>
<td>25</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Not smoked for 1 - 10 years</td>
<td>62</td>
<td>4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Recent quitter</td>
<td>462</td>
<td>31</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Current smoker</td>
<td>157</td>
<td>11</td>
<td>14</td>
<td>6</td>
</tr>
<tr>
<td>Mean age in years (SD)</td>
<td></td>
<td></td>
<td>60.6 (10.4)</td>
<td>60</td>
</tr>
<tr>
<td>Mean CRF in mL min⁻¹ kg⁻¹ (SD)</td>
<td>20.5 (8.5)</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Mean HADS-D (SD)</td>
<td>4.2 (3.3)</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Mean HADS-A (SD)</td>
<td>6.5 (4)</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Mean BMI in Kg m⁻² (SD)</td>
<td>27.5 (4.3)</td>
<td>-</td>
<td>28</td>
<td></td>
</tr>
</tbody>
</table>

Sample characteristics (n = 1489). MI myocardial infarction, CABG coronary artery bypass grafting, PCI percutaneous coronary intervention, CRF cardiorespiratory fitness, HADS-D depression subscale of hospital anxiety and depression scale, HADS-A anxiety subscale of hospital anxiety and depression scale, BMI body mass index. HADS subscales score: 0-7 Normal, 8-10 Mild, 11-15 Moderate, 16-21 Severe.
<table>
<thead>
<tr>
<th></th>
<th>Interaction</th>
<th>Depression</th>
<th>Sex</th>
<th>Age</th>
<th>Diagnosis</th>
<th>Diabetes</th>
<th>Smoking</th>
<th>BMI</th>
<th>Anxiety</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model A (Sex¹)</td>
<td>-0.3**</td>
<td>-0.25*</td>
<td>5.56**</td>
<td>-0.35**</td>
<td>0.33**</td>
<td>-2.66**</td>
<td>-0.36*</td>
<td>0.45**</td>
<td>0.04</td>
<td>0.445</td>
</tr>
<tr>
<td>Model B (BMI)</td>
<td>0.05**</td>
<td>-0.48**</td>
<td>4.05**</td>
<td>-0.34**</td>
<td>0.31**</td>
<td>-2.55**</td>
<td>-0.37**</td>
<td>0.46**</td>
<td>0.05</td>
<td>0.45</td>
</tr>
<tr>
<td>Model C (Sex¹ + BMI)</td>
<td>Sex: -0.44**</td>
<td>-0.49**</td>
<td>4.33**</td>
<td>-0.34**</td>
<td>0.31**</td>
<td>-2.51**</td>
<td>-0.36**</td>
<td>0.46**</td>
<td>0.05</td>
<td>0.454</td>
</tr>
</tbody>
</table>

Unstandardized regression coefficients in statistically significant moderations. ¹ Sex coding: 0 for female; 1 for male, BMI body mass index, * p < 0.05, ** p < 0.025.
Figure 4.1

Moderation of the relationship between depression and CRF by BMI in male patients (n = 1198). CRF cardiorespiratory fitness, HADS-D depression subscale of hospital anxiety and depression scale, BMI body mass index.

Figure 4.2

Moderation of the relationship between depression and CRF by BMI in female patients (n = 291). CRF cardiorespiratory fitness, HADS-D depression subscale of hospital anxiety and depression scale, BMI body mass index.
Discussion

Depression symptom severity was found to be inversely related to CRF in this study. This finding is consistent to the findings of our previous meta-analyses, which show an inverse relationship between depression and CRF in healthy individuals (Papasavvas et al., 2016) and patients with heart disease (Papasavvas et al., 2017). A combined moderation of this relationship by sex and BMI was also found, with low BMI enhancing the relationship in male patients and rendering it statistically significant in female patients and high BMI attenuating the relationship in male patients and rendering it statistically not significant in female patients with coronary heart disease.

The blunting effect of high BMI [BMI > 26.8 kg · m\(^{-2}\) (sample median)] in the relationship between depression symptom severity and CRF may be explained by the fact that in this group there was a positive correlation between BMI and CRF (r = 0.22, p < 0.001) and no negative correlation between BMI and depression symptom severity (r = -0.04, p = 0.33). Respectively, the enhancing effect of low BMI (BMI < 26.8 kg · m\(^{-2}\)) in the relationship between depression symptom severity and CRF may be explained by the fact that in that group there was a positive correlation between BMI and CRF (r = 0.26, p < 0.001) and a negative correlation between BMI and depression symptom severity (r = -0.14, p < 0.001).

The positive correlation between BMI and CRF in this study might seem unexpected. However, although BMI and CRF may be inversely related in healthy individuals (Farrell et al., 2013, Erez et al., 2015, Chen et al., 2017, Pandey et al., 2017), the relationship seems to be more complex in patients with heart disease. In a large group of men (n = 9563) with known or suspected coronary heart disease BMI was indeed negatively related to CRF (McAuley et al., 2012). However, in another large group of men (n = 12147) with known or suspected coronary heart disease CRF was the same in patients with normal weight (BMI 18.5 – 24.9 kg · m\(^{2}\)) and overweight patients (BMI 25.0 – 29.9 kg · m\(^{2}\)), and higher than the CRF of underweight patients (BMI < 18.5 kg · m\(^{2}\)) (McAuley et al., 2010). Additionally, in patients with heart disease (n = 2324) only obese III patients
(BMI ≥ 40 kg · m²), not overweight (BMI 25.0 – 29.9 kg · m²), obese I (BMI 30 – 34.9 kg · m²), or obese II (BMI 35 – 39.9 kg · m²), had lower CRF than patients with normal weight (Horwich et al., 2009). It should be noted that obese III patients comprised only 1% of our sample.

The inverse association between depression symptom severity and CRF may be explained via behavioral pathways. A plausible pathway is that depressed patients may not be motivated to engage in physical activity and therefore may have lower fitness. There is evidence that depression is associated with concurrent physical inactivity (Harris et al., 2006, Lysy et al., 2008, Mc Dowell et al., 2018) and that depression may also predict physical activity levels (Roshanaei-Moghaddam et al., 2009, Alosco et al., 2012). Bidirectional pathways between depression and CRF may also be present (Azevedo Da Silva et al., 2012), as physical activity has also been reported to protect against the development of depression (Schuch et al., 2018), potentially indicating the existence of biological pathways. It should be noted that physical activity was not assessed in this study and therefore, its potential moderating effect could not be tested. More research is warranted in this area.

The inverse association between depression symptom severity and CRF moderated by sex and BMI in patients with heart disease has implications regarding patient assessment and clinical management. Clinicians may consider that patients with low CRF may exhibit elevated depressive symptoms, especially if they are men and have normal weight or are underweight. In such cases a more thorough assessment of depressive symptoms may be warranted and clinicians might be prepared to face depression-related treatment challenges, such as adherence to medication (Gehi et al., 2005, Goldstein et al., 2017), and enrollment (Kotseva et al., 2018) and adherence (Casey et al., 2008, McGrady et al., 2009, Swardfager et al., 2011) to cardiac rehabilitation. Respectively, if patients exhibit elevated depressive symptoms they may have low CRF. In such cases clinicians may consider a more thorough assessment of CRF (i.e. a symptom-limited exercise test) and focus on exercise prescription to increase CRF (Ismail et al., 2013, Hollings et al., 2017, Hannan et al., 2018), a strong and independent predictor of mortality (Kavanagh et al., 2002, Martin et al., 2013).
The inverse association between depression symptom severity and CRF also has implications regarding research. Regardless if the association is a result of unidirectional causation, bidirectional causation, or the influence of other variable(s), it warrants further research on the effects of treating the one on the other. Improving depression might increase CRF and, respectively, increasing CRF might improve depression. Exercise-based cardiac rehabilitation that improves CRF (Haykowsky et al., 2013, Elliott et al., 2015, Santos et al., 2018) has also been reported to reduce depressive symptoms (Blumenthal et al., 2012, Tu et al., 2014); however, more research is required to determine the optimum frequency, intensity, type, and duration of exercise for reducing depressive symptoms. There is also evidence that psychological interventions reduce depressive symptoms in patients with heart disease (Richards et al., 2017) but their effects on CRF have not been studied yet, to my knowledge.

Limitations

Patients who were referred to cardiac rehabilitation but were not enrolled or were enrolled without initial assessment of fitness were excluded from this study. These patients may typically form a significant part of patients referred to cardiac rehabilitation; therefore, their exclusion may affect the generalizability of the results. Patients who underwent initial assessment without complete data for all studied variables were also excluded. There was no record of physical activity and dietary factors; therefore, their impact was not studied. Depressive symptoms were measured using the HADS. Although the HADS has been reported to have acceptable properties for detecting depression in patients with heart disease (Stafford et al., 2007, Haddad et al., 2013), it lacks items related to somatic symptoms of depression, such as fatigue and sleep disturbance, and it was not developed specifically for patients with heart disease. Perhaps the Cardiac Depression Scale (Hare and Davis 1996) might be more appropriate to assess depression in patients with heart disease, as this scale is comprehensive and was developed specifically for this patient population.
It should also be noted that this study used data that were collected between 1993 and 2002. As it can be seen in Table 4.1, there are differences in variables that may affect depression and CRF between that period and 2017, which reflect the evolution of the management of heart disease and the utilization of cardiac rehabilitation throughout these years. The mean age of patients enrolled in cardiac rehabilitation has increased from 61 to 67 years, female patients enrollement has increased from 20% to 29%, primary PCI has increased from 5% to 29%, and concurrently, medical treatment of myocardial infarction has decreased from 56% to 16%. Age, sex, and management of heart disease can affect depressive symptoms and CRF and therefore, generalizing the findings of this study on the current population of patients entering cardiac rehabilitation should be done with caution.

Conclusions

An inverse relationship between depression symptom severity and CRF that was moderated by sex and BMI was found in this study. Lower BMI enhanced the relationship in both sexes and higher BMI attenuated the relationship in male and rendered it statistically not significant in female patients with heart disease. The findings have implications regarding the assessment and clinical management of patients with heart disease, and warrant further research on the interaction between depression and CRF in these patients.


Chapter 5

Translation and validation of the Cardiac Depression Scale to Arabic.
Translation and validation of the Cardiac Depression Scale to Arabic (Papasavvas et al., 2016)

T. Papasavvas\textsuperscript{a,b}, H. Al-Amin\textsuperscript{c,d}, H.F. Ghabrash\textsuperscript{e} and D. Micklewright\textsuperscript{b}

Abstract

Background: The Cardiac Depression Scale (CDS) has been designed to measure depressive symptoms in patients with heart disease. There is no Arabic version of the CDS. We translated and validated the CDS in an Arabic sample of patients with heart disease.

Methods: Forward and back translation of the CDS was followed by assessment of cultural relevance and content validity. The Arabic version of the CDS (A-CDS) and the Arabic version of the Hospital Anxiety and Depression Scale (A-HADS) were then administered to 260 Arab in-patients with heart disease from 18 Arabic countries. Construct validity was assessed using exploratory factor analysis with polychoric correlations. Internal consistency was assessed using ordinal reliability alpha and item-to-factor polychoric correlations. Concurrent validity was assessed using Pearson’s correlation coefficient between the A-CDS and the depression subscale of the A-HADS (A-HADS-D).

Results: Cultural relevance and content validity of the A-CDS were satisfactory. Exploratory factor analysis revealed three robust factors, without cross-loadings, that formed a single dimension. Internal consistency was high (ordinal reliability alpha for the total scale and the three factors were .94, .91, .86, and .87, respectively; item-to-factor correlations ranged from .77 to .91). Concurrent validity was high ($r = .72$). The A-CDS demonstrated a closer to normal distribution of scores than the A-HADS-D.

\textsuperscript{a}Cardiac Rehabilitation Department, Heart Hospital, Hamad Medical Corporation, Doha, Qatar  
\textsuperscript{b}School of Biological Sciences, University of Essex, Colchester, UK  
\textsuperscript{c}Weill Cornell Medical College in Qatar, Doha, Qatar  
\textsuperscript{d}Psychiatry Department, Rumeilah Hospital, Hamad Medical Corporation, Doha, Qatar  
\textsuperscript{e}The Northern Hospital, Northwestern Mental Health, Melbourne Health, Melbourne, Australia
Limitations: Sensitivity and specificity of the A-CDS were not objectively assessed.

Conclusions: The A-CDS appears to be a valid and reliable instrument to measure depressive symptoms in a representative sample of Arab in-patients with heart disease.

1. Introduction

Depression is a leading cause of disability worldwide, afflicting more than 350 million people of all ages (WHO 2012). Its prevalence ranges from 6.6% to 21% across countries (Kessler and Bromet 2013), rising to 18.3% in Qatar (Bener et al., 2015). Nearly 20% of patients with heart disease suffer from depression (Thombs et al., 2006, Carney and Freedland 2008, Elderon and Whooley 2013), while depression following myocardial infarction is associated with 2.25- to 2.38-fold risk of all-cause mortality and 2.59- to 2.71-fold risk of cardiac mortality (van Melle et al., 2004, Meijer et al., 2011). In patients with heart failure, depression is associated with 51% increased risk of all-cause mortality and 119% increased risk of cardiac mortality (Fan et al., 2014). Based on the consistency of evidence relating depression to adverse outcomes after acute coronary syndrome, the American Heart Association has elevated depression to the status of a risk factor for poor prognosis in patients with acute coronary syndrome (Lichtman et al., 2014).

Given the impact of depression on the prognosis of patients with heart disease, it is not surprising that screening for depressive symptoms in patients with coronary heart disease has also been recommended (Lichtman et al., 2008). Several psychometric scales have been used for this purpose, including the Beck Depression Inventory II (BDI-II) (Beck et al., 1996), the Hospital Anxiety and Depression Scale (HADS) (Zigmond and Snaith 1983), the Hamilton Rating Scale for Depression (HAM-D) (Hamilton 1960), and the Centre for Epidemiologic Studies Depression Scale (CES-D) (Radloff 1977). These scales are validated and exhibit satisfactory internal consistency [$\alpha = .80$ to $.90$ for BDI-II (Beck et al., 1996), $\alpha = .67$ to $.90$ for the depression subscale of HADS (Bjelland et al., 2002), $\alpha = .77$ to $.81$ for HAM-D (Trajkovic et al., 2011), and $\alpha = .85$ to $.90$ for CES-D (Radloff 1977)]. However, none of these scales were developed specifically for patients with heart
disease and therefore, their psychometric properties may not apply to this patient group. Moreover, they may not be comprehensive enough to detect depression in patients with heart disease. The HADS, for example, lacks items related to somatic symptoms of depression, such as fatigue and sleep disturbance. Furthermore, they may not be sensitive enough to detect minor depression, which is clinically significant (Bush et al., 2001, Catipovic-Veselica et al., 2007, Lossnitzer et al., 2013). The BDI-II, for example, has a positively skewed distribution of scores which results in low scores clustering and poor differentiation (Di Benedetto et al., 2006). It is clear that generic scales have significant limitations in assessing depression and depressive symptoms in patients with heart disease (Vieweg et al., 2011).

The Cardiac Depression Scale (CDS) (Hare and Davis 1996) is the only psychometric scale designed to measure depressive symptoms in patients with heart disease. It was validated in Australian outpatients of a cardiology clinic comprising a wide range of diagnosis including angina, heart failure, post-myocardial infarction, post-surgery, valve disease and arrhythmias. The CDS consists of 26 items scored on a seven-point Likert scale ranging from strongly disagree (1) to strongly agree (7) and it exhibited satisfactory correlations with clinical assessment (r = .67) and the BDI (r = .73), as well as satisfactory internal reliability (α = .90) (Hare and Davis 1996). It also demonstrated a normal distribution of scores compared to the strongly positively skewed distribution of the BDI (Hare and Davis 1996), which enables CDS to differentiate low scores and therefore be sensitive enough to detect minor depression. These results have been replicated in other English-speaking samples (Birks et al., 2004, Wise et al., 2006, Kiropoulos et al., 2012, Ski et al., 2012).

The CDS has been translated and validated in German (Hare et al., 2000), Chinese (Wang et al., 2008), and Iranian (Gholizadeh et al., 2010) patients with heart disease. There is no Arabic version of the CDS; therefore, the purpose of this study was to develop an Arabic translation of the CDS and validate it in Arab patients with heart disease.
2. Methods

This study was conducted in two phases:

1. Translation of the original version of the CDS to Arabic ensuring cultural relevance and content validity.

2. Evaluation of construct validity, internal consistency, and concurrent validity of the Arabic version of the CDS (A-CDS).

2.1 Translation of the Cardiac Depression Scale to Arabic

Forward translation (English to Arabic) was conducted by two bilingual Arab experts (psychiatrist and cardiac rehabilitation specialist with Master degree). After translating the scale separately, they met and agreed on the final translation of each item. Back translation (Arabic to English) was conducted by a certified translation services company. Subsequently, the principal investigator (T.P.), the two bilingual Arab experts and two Arab representatives from the translation services company met and finalized each item of the A-CDS. There was unanimous agreement on conceptual equivalence in every item.

2.2 Cultural relevance and content validity

The cultural relevance and the content validity of the A-CDS were evaluated by a panel of six bilingual Arab clinicians (three psychiatrists, one consultant cardiologist, one physiatrist, and one nurse specialized in quality improvement). They were asked to rate the cultural relevance and the content validity of each item by using a 4-point Likert scale: 1 = not relevant, 2 = somewhat relevant, 3 = quite relevant, and 4 = highly relevant. The scale content validity index (S-CVI) was calculated as the proportion of items that achieved a rating of 3 or 4 by all clinicians. An S-CVI score of 80% or higher is indicative of satisfactory content validity (Davis 1992).
2.3 Evaluation of construct validity, internal consistency, and concurrent validity

2.3.1 Instruments

2.3.1.1 Cardiac Depression Scale. The CDS is a self-administered depression scale for use in patients with heart disease. It consists of 26 items scored on a seven-point Likert scale ranging from strongly disagree (1) to strongly agree (7). In the original study (Hare and Davis 1996), seven factors have been reported to comprise the scale: Sleep, anhedonia, uncertainty, mood, cognition, hopelessness, and inactivity. These factors formed two dimensions in a second-order factor analysis. The scale exhibited satisfactory internal consistency (\( \alpha = .90 \)) and correlated with the BDI and clinical assessment (\( r = .73 \) and \( r = .67 \), respectively).

2.3.1.2 Hospital Anxiety and Depression Scale. The HADS is a self-administered depression scale developed in a hospital medical outpatient clinic (Zigmond and Snaith 1983). It consists of 14 items scored on a four-point Likert scale (0-3) that are evenly divided into an anxiety subscale and a depression subscale. The scale comprises two factors (anxiety and depression) and it exhibits satisfactory internal consistency (\( \alpha = .86 \) for the total scale, \( \alpha = .77 \) for the anxiety subscale, and \( \alpha = .82 \) for the depression subscale) (Zigmond and Snaith 1983). An Arabic version of the HADS has been developed (el-Rufaie and Absood 1987, el-Rufaie and Absood 1995) with satisfactory correlation with clinical evaluation (\( r = .82 \)) and internal consistency (\( \alpha = .88 \)).

2.3.2 Procedure

A subject to item ratio of 10:1 was used to estimate the sample size, as it is often used in the absence of clear scientifically sound recommendations on this topic (Anthoine et al., 2014). A convenience sample of 260 adult (18 years and above) patients admitted to the Heart Hospital, a tertiary care hospital for patients with heart disease in Doha and member of Hamad Medical Corporation, were recruited from June 2014 to February 2015 based on data collectors availability (see patient flow in Fig. 5.1). Inclusion criteria were the following:
Figure 5.1

Patient flow during recruitment period.
(a) having a diagnosis of heart disease, (b) not suffering from a major psychiatric disease, including schizophrenia, bipolar disorder, and dementia (this was confirmed by reviewing the medical records and asking the participant and family members), (c) being a national of a country where Arabic is an official language, and (d) having Arabic as mother tongue and preferred mode of oral and written communication. The study was approved by the ethics committee of Hamad Medical Corporation. Nurses specialized in cardiac rehabilitation explained the study to eligible participants, provided a sheet with study’s details, answered any questions, and obtained verbal consent from participants. Participants were given an envelope with the A-CDS and the Arabic version of the HADS (A-HADS) (el-Rufaie and Absood 1987, el-Rufaie and Absood 1995) and were left alone to complete them and seal the envelope. The sealed envelopes were collected later within the day by the same nurse who provided them. Demographic and clinical data were obtained from medical records. All administering personnel were blinded to outcomes and interpretation.

2.3.3 Data analysis

Likert-type item-level data are rarely continuous and normally distributed (Bernstein and Teng 1989, Jamieson 2004) and our data were not an exclusion, either. Instead of Pearson’s correlations, polychoric correlations are recommended to analyse Likert-type data (Basto and Pereira 2012, Gaskin and Happell 2014); therefore, construct validity was evaluated using exploratory factor analysis with polychoric correlations. The Kaiser – Meyer – Olkin (KMO) test for the total scale and each item was used to assess sampling adequacy. A bare minimum of .50 is required for the sample to be adequate (Kaiser 1974). Parallel analysis (Horn 1965), often recommended as one of the best methods to determine the number of factors to retain (Zwick and Velicer 1986, Hayton et al., 2004, Peres-Neto et al., 2005), using principal component analysis as the method for extraction with the original data randomized (permutation data) and the mean eigenvalue criterion (Garrido et al., 2013) was used to determine the number of factors to retain. Ordinary least squares factor analysis (equivalent to minimum residuals (Harman 1960), also known as unweighted least squares factor analysis (Lee et al., 2012)) was used to extract data, as it is suggested to be the most appropriate extraction method.
for ordinal data (Gaskin and Happell 2014). Oblique rotation (quartimin) was used to simplify the structure because psychological factors are typically correlated, and this was the case in our data, as well. A factor loading of .40 or greater was set as a cut-off point for significant loading. The quality of the factorial model was assessed using the goodness of fit index, which can be defined as the fraction of the correlations of the observed variables that are explained by the model. Values greater than .90 usually indicate a good fit.

Internal consistency was evaluated using the ordinal version of Cronbach’s alpha, called ordinal reliability alpha (Zumbo et al., 2007), and by estimating the item-to-factor polychoric correlation coefficient, which is the correlation coefficient between the item and its factor with the item removed. Concurrent validity was evaluated using Pearson’s correlation coefficient between the A-CDS (and the extracted factors) and the depression subscale of the Arabic version of the HADS (A-HADS-D) (el-Rufaie and Absood 1987, el-Rufaie and Absood 1995). The A-HADS was selected instead of the Arabic version of the HAM-D (Hamdi et al., 1997) because it has demonstrated higher internal consistency. Arabic versions of the BDI-II (Abdel-Khalek 1998, Al-Musawi 2001, Al-Turkait and Ohaeri 2010) and the CES-D (Ghubash et al., 2000, Kazarian and Taher 2010) were not selected because they have been validated in college students and community samples, not in clinical populations.

If one item of the scale was left blank the average item score was inputted. If more than one items of the scale were left blank the scale was discarded. The SPSS ver. 22 with an R-Menu for ordinal factor analysis (Basto and Pereira 2012) was used for statistical analysis.

3. Results

3.1 Cultural relevance and content validity

The expert panel evaluated the scale and gave an S-CVI score of 69% and 81% for cultural relevance and content validity, respectively, along with comments and suggestions. Two bilingual Arab psychiatrists (investigators H.A and H.G.) revised the items and sent the revised scale back to the expert panel for a second
evaluation. The expert panel evaluated the revised scale and gave an S-CVI score of 96% for both cultural relevance and content validity. The item that did not achieve a rating of 3 or 4 by all experts was item 10 “I feel like I’m living on borrowed time”.

3.2 Sample characteristics

Demographic and clinical characteristics of the sample can be seen in Table 5.1. The mean age was 54 (S.D. = 12.7, range = 19 – 88) years. The sample included nationals from 18 Arab countries and the majority was male (85.4%) and married (88.8%) participants. The mean A-CDS score was 82.5 (S.D. = 26.7, median = 78, range = 28 – 153). Using a cut-off score of 100 for severe depression and 90 for mild to moderate depression (Wise, Harris et al. 2006), 61 patients (23.5%) demonstrated severe depression and additional 30 patients (11.5%) demonstrated mild to moderate depression.
<table>
<thead>
<tr>
<th>Variable</th>
<th>n (N = 260)</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>222</td>
<td>85.4</td>
</tr>
<tr>
<td>Female</td>
<td>38</td>
<td>14.6</td>
</tr>
<tr>
<td><strong>Nationality</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qatar</td>
<td>74</td>
<td>28.5</td>
</tr>
<tr>
<td>Egypt</td>
<td>67</td>
<td>25.8</td>
</tr>
<tr>
<td>Syria</td>
<td>21</td>
<td>8.1</td>
</tr>
<tr>
<td>Palestinian Territory</td>
<td>20</td>
<td>7.7</td>
</tr>
<tr>
<td>Sudan</td>
<td>19</td>
<td>7.3</td>
</tr>
<tr>
<td>Jordan</td>
<td>15</td>
<td>5.8</td>
</tr>
<tr>
<td>Yemen</td>
<td>10</td>
<td>3.8</td>
</tr>
<tr>
<td>Lebanon</td>
<td>7</td>
<td>2.7</td>
</tr>
<tr>
<td>Oman</td>
<td>5</td>
<td>1.9</td>
</tr>
<tr>
<td>Iraq</td>
<td>5</td>
<td>1.9</td>
</tr>
<tr>
<td>Bahrain</td>
<td>4</td>
<td>1.5</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>3</td>
<td>1.2</td>
</tr>
<tr>
<td>Somalia</td>
<td>2</td>
<td>0.8</td>
</tr>
<tr>
<td>Tunisia</td>
<td>2</td>
<td>0.8</td>
</tr>
<tr>
<td>Algeria</td>
<td>2</td>
<td>0.8</td>
</tr>
<tr>
<td>Mauritania</td>
<td>2</td>
<td>0.8</td>
</tr>
<tr>
<td>Eritrea</td>
<td>1</td>
<td>0.4</td>
</tr>
<tr>
<td>Kuwait</td>
<td>1</td>
<td>0.4</td>
</tr>
<tr>
<td><strong>Marital status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>231</td>
<td>88.8</td>
</tr>
<tr>
<td>Single</td>
<td>19</td>
<td>7.3</td>
</tr>
<tr>
<td>Divorced</td>
<td>5</td>
<td>1.9</td>
</tr>
<tr>
<td>Widow(er)</td>
<td>5</td>
<td>1.9</td>
</tr>
<tr>
<td><strong>Diagnosis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute coronary syndrome</td>
<td>130</td>
<td>50.0</td>
</tr>
<tr>
<td>Coronary artery disease</td>
<td>59</td>
<td>22.7</td>
</tr>
<tr>
<td>Heart failure</td>
<td>28</td>
<td>10.8</td>
</tr>
<tr>
<td>Arrhythmia</td>
<td>24</td>
<td>9.2</td>
</tr>
<tr>
<td>Valve disease</td>
<td>9</td>
<td>3.5</td>
</tr>
<tr>
<td>Cardiomyopathy</td>
<td>5</td>
<td>1.9</td>
</tr>
<tr>
<td>Aortic dissection</td>
<td>2</td>
<td>0.8</td>
</tr>
<tr>
<td>Pericarditis</td>
<td>2</td>
<td>0.8</td>
</tr>
<tr>
<td>Myopericarditis</td>
<td>1</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Demographic and clinical characteristics of the sample.
3.3 Construct validity

No scale had more than one item left blank. One item had been left blank in five A-CDS scales (different item in each scale), in which the average item score was inputted. Additionally, four patients ticked more than one box in a single item in the A-CDS (different item in each patient); these patients were asked to clarify their response by the same nurse who administered the scales at the same day. The KMO sampling adequacy test for the total scale was .89. Measures of sampling adequacy in the items ranged from .75 to .95. Item skewnesses ranged from -.07 to 2.13 and kurtoses from -1.58 to 3.95. All items had statistically significant Kolmogorov-Smirnov and Shapiro-Wilk tests (p < .001) and visual inspection confirmed significant deviations from normal distribution in all items. Parallel analysis using principal component analysis as the method of extraction with the original data randomized (permutation data) and the mean eigenvalue criterion indicated that three factors should be retained. Ordinary least squares factor analysis with oblique (quartimin) rotation created a three-factor model accounting for 48% of variance. Factor 1 included the items from the original scale (Hare and Davis 1996) factors “Sleep”, “Uncertainty”, “Mood”, and “Inactivity”. Factor 2 included the items from factors “Anhedonia” and “Cognition”, and Factor 3 included the items from factor “Hopelessness”. Loadings ranged from .40 to .84 and there were no cross-loadings with loadings greater than .40 (see Table 5.2). All factors were correlated with correlations ranging from .37 to .56. The goodness of fit index was .92. Using the same methodology that was used in the first-order factor analysis, a second-order factor analysis produced a single dimension containing all original three factors with loadings ranging from .67 (Factor 3) to .87 (Factor 1) and accounting for 58% of variance.

3.4 Internal consistency

Item-to-factor correlation coefficients ranged from .77 to .91. Ordinal reliability alpha for Factor 1, 2, and 3 were .91, .86, and .87, respectively. The ordinal reliability alpha for the total scale was .94.
Table 5.2

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
<th>Factor loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>I wake up in the early hours of the morning and cannot get back to sleep</td>
<td>.76</td>
</tr>
<tr>
<td>5</td>
<td>I am concerned about the uncertainty of my health</td>
<td>.75</td>
</tr>
<tr>
<td>6</td>
<td>I may not recover properly</td>
<td>.74</td>
</tr>
<tr>
<td>7</td>
<td>My sleep is restless and disturbed</td>
<td>.74</td>
</tr>
<tr>
<td>1</td>
<td>I have dropped many of my interests and activities</td>
<td>.62</td>
</tr>
<tr>
<td>18</td>
<td>Things which I regret about my life are bothering me</td>
<td>.62</td>
</tr>
<tr>
<td>8</td>
<td>I am not the person I used to be</td>
<td>.62</td>
</tr>
<tr>
<td>24</td>
<td>I lose my temper more easily nowadays</td>
<td>.56</td>
</tr>
<tr>
<td>22</td>
<td>I seem to get more easily irritated by others than before</td>
<td>.53</td>
</tr>
<tr>
<td>3</td>
<td>I can’t be bothered doing anything much</td>
<td>.47</td>
</tr>
<tr>
<td>17</td>
<td>My problems are not yet over</td>
<td>.46</td>
</tr>
<tr>
<td>21</td>
<td>I become tearful more easily than before</td>
<td>.46</td>
</tr>
<tr>
<td>13</td>
<td>The possibility of sudden death worries me</td>
<td>.45</td>
</tr>
<tr>
<td>25</td>
<td>I feel frustrated</td>
<td>.44</td>
</tr>
<tr>
<td>26</td>
<td>I am concerned about my capacity for sexual activity</td>
<td>.42</td>
</tr>
<tr>
<td>16</td>
<td>I get hardly anything done</td>
<td>.40</td>
</tr>
<tr>
<td>20</td>
<td>My memory is as good as it always was</td>
<td>.84</td>
</tr>
<tr>
<td>15</td>
<td>My mind is as fast and alert as always</td>
<td>.82</td>
</tr>
<tr>
<td>19</td>
<td>I gain just as much pleasure from my leisure activities as I used to</td>
<td>.61</td>
</tr>
<tr>
<td>2</td>
<td>My concentration is as good as it ever was</td>
<td>.60</td>
</tr>
<tr>
<td>23</td>
<td>I feel independent and in control of my life</td>
<td>.57</td>
</tr>
<tr>
<td>4</td>
<td>I get pleasure from life at present</td>
<td>.46</td>
</tr>
<tr>
<td>12</td>
<td>I feel in good spirits</td>
<td>.43</td>
</tr>
<tr>
<td>11</td>
<td>Dying is the best solution for me</td>
<td>.80</td>
</tr>
<tr>
<td>14</td>
<td>There is only misery in the future for me</td>
<td>.77</td>
</tr>
<tr>
<td>10</td>
<td>I feel like I’m living on borrowed time</td>
<td>.68</td>
</tr>
</tbody>
</table>

Factor loadings for the Arabic version of the Cardiac Depression Scale using ordinary least squares factor analysis with oblique (quartimin) rotation.

3.5 Concurrent validity

The A-CDS correlated significantly with the depression and anxiety subscales of the A-HADS ($r = .72$ and $r = .78$, respectively, $p < .001$ for both). All factors of the A-CDS also correlated significantly with both subscales of the HADS (see Table 5.3). The distribution of scores of the A-CDS had skewness .48 and kurtosis -.36 and
was closer to normal (see Fig. 5.2) than the distribution of scores of the depression subscale of the A-HADS, which had skewness .96 and kurtosis .92 (see Fig. 5.3).

### Table 5.3

<table>
<thead>
<tr>
<th>A-CDS</th>
<th>A-HADS Depression</th>
<th>A-HADS Anxiety</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1 (sleep, uncertainty, mood, inactivity)</td>
<td>.66</td>
<td>.77</td>
</tr>
<tr>
<td>Factor 2 (anhedonia, cognition)</td>
<td>.60</td>
<td>.53</td>
</tr>
<tr>
<td>Factor 3 (Hopelessness)</td>
<td>.52</td>
<td>.55</td>
</tr>
<tr>
<td>Total scale</td>
<td>.72</td>
<td>.78</td>
</tr>
</tbody>
</table>

Pearson’s correlation coefficients between the A-CDS and the A-HADS domains. All correlations are significant at .001.

*A-CDS*, Arabic version of the Cardiac Depression Scale; *A-HADS*, Arabic version of the Hospital Anxiety and Depression Scale.
Figure 5.2

Distribution of scores of the Arabic version of the Cardiac Depression Scale. A-CDS, Arabic version of the Cardiac Depression Scale.

Figure 5.3

Distribution of scores of the depression subscale of the Arabic version of the Hospital Anxiety and Depression Scale. A-HADS, Arabic version of the Hospital Anxiety and Depression Scale.
4. Discussion

This study was the first attempt to translate the CDS to Arabic and validate it in an in-patient population with heart disease. The findings demonstrated that the A-CDS is a valid and reliable instrument for measuring depressive symptoms in Arab patients with heart disease. The population was multinational comprising nationals from 18 Arabic countries and included a wide range of age and a typical spectrum of heart disease diagnoses; therefore, the generalizability of the results could benefit the vast majority of Arab patients with heart disease.

The CDS is most commonly used for depression screening in patients with heart disease. Using recommended cut-off scores (Wise et al., 2006), the A-CDS demonstrated that 23.5% of the study population exhibited major depression. This is in line with international data (Thombs et al., 2006, Carney and Freedland 2008, Elderon and Whooley 2013) and adds to the validity of the A-CDS. This finding, however, should be interpreted with caution because these cut-off scores have been produced using the Geriatric Depression Scale—Short Form (GDS-SF) (Sheikh and Yesavage 1986) as a criterion measure. The GDS-SF has been validated in a wide variety of populations, including primary care clinics, geriatric outpatient clinics, nursing homes, and medical and neurological in-patients (Wancata et al., 2006), but not specifically in patients with heart disease; therefore, its psychometric properties in patients with heart disease are unknown.

The A-CDS demonstrated a closer to normal distribution of scores compared to the positively skewed distribution of scores of the A-HADS-D. Similar findings have been reported when comparing the CDS with the BDI (Hare and Davis 1996, Birks et al., 2004, Di Benedetto et al., 2006, Gholizadeh et al., 2010). This can be explained by the fact that the CDS was originally developed to assess “adjustment disorder with depressed mood” rather than major depressive disorder (Hare and Davis 1996). A close to normal distribution of scores enables the A-CDS to detect and differentiate between mild to moderate depressive symptoms, which have clinical and prognostic significance (Bush et al., 2001, Catipovic-Veselica et al., 2007, Lossnitzer et al., 2013);
therefore, using the A-CDS to screen Arab patients with heart disease for depression can provide clinicians with better information in order to decide whether the patient needs to undergo further psychological evaluation or not; and using the A-CDS to assess changes in depressive symptoms could produce more accurate results, especially within the mild to moderate depressive symptoms spectrum.

The exploratory factor analysis created a stable, parsimonious, and meaningful three-factor model. These factors formed a single dimension in a second-order analysis. Sample adequacy and goodness of fit were satisfactory. In the original study (Hare and Davis 1996) a seven-factor model was reported, which formed two dimensions in a second-order analysis. Since then three exploratory factor analyses have been performed in the English scale (Birks et al., 2004, Wise et al., 2006, Kiropoulos et al., 2012), one in a German version (Hare et al., 2000), one in a Chinese version (Wang et al., 2008), and one in an Iranian version (Gholizadeh et al., 2010). The resulted factorial models differed: One-factor (Birks et al., 2004), two-factor (Gholizadeh et al., 2010), six-factor (Wise et al., 2006, Wang et al., 2008, Kiropoulos et al., 2012), and seven-factor (Hare et al., 2000) models were reported. The differences in the models, including our model, can be attributed to methodological differences between studies, including sample size, sample language and culture, heart disease diagnosis, and method of analysis. We could not compare sample adequacy and goodness of fit measures because they were not reported in any of the studies. Furthermore, none of the studies used polychoric correlations, although Likert-type item-level data are rarely continuous and normally distributed (Bernstein and Teng 1989, Jamieson 2004). Regarding determining the number of factors to extract, most studies aimed to replicate the methodology of the original study (Hare and Davis 1996) and used the eigenvalues greater than one rule (Kaiser 1960). However, there is agreement in the literature that this criterion is one of the least accurate criteria to determine the number of factors to extract and that it usually overestimates this number (Velicer and Jackson 1990, Fabrigar et al., 1999, Costello and Osborne 2005, Lance et al., 2006). Parallel analysis, on the other hand, is considered one of the best methods for this task (Zwick and Velicer 1986, Hayton et al., 2004, Peres-Neto et al., 2005). Last, different methods were used to extract
data, including maximum likelihood analysis (Hare et al., 2000), principal components analysis (Wise et al., 2006, Wang et al., 2008, Kiropoulos et al., 2012) and principal axis factoring (Birks et al., 2004, Gholizadeh et al., 2010). It should be noted that for ordinal data the recommended method for data extraction is ordinary least squares factor analysis (Gaskin and Happell 2014).

The total A-CDS and the three comprising factors exhibited excellent internal consistency ranging from .87 to .94. This was also evident by the high item-to-factor correlations. Moreover, the three factors correlated moderately to each other indicating a balance between homogeneity and differentiation of the scale.

The moderate to high correlation between the A-CDS and the A-HADS-D indicated satisfactory concurrent validity. It should be noted, though, that the criterion measure (HADS and A-HADS) has been developed and validated in hospital medical clinic outpatients (Zigmond and Snaith 1983), and primary healthcare outpatients (el-Rufai and Absood 1987, el-Rufai and Absood 1995), not in patients with heart disease. Moreover, it lacks items related to somatic symptoms of depression. Therefore, although it may be a good choice among available scales, it is not the best criterion measure for concurrent validity in this study; a clinical evaluation would be the best criterion measure but it was not feasible in our study. Therefore, the sensitivity and specificity of the A-CDS could not be objectively assessed.

We also found a moderate to high correlation between the A-CDS and the anxiety subscale of the A-HADS, which was stronger than the correlation between the A-CDS and A-HADS-D (.78 vs .72, respectively). Although the difference in the correlations may not be clinically significant, it suggests that there was a comorbidity of depression and anxiety in our sample. This may have two not mutually exclusive explanations: (a) The comorbidity of depression and anxiety is evident worldwide (Kessler et al., 2015) and in the Arabic population specifically (Belzer and Schneier 2004, Ohaeri et al., 2010, Al-Turkait et al., 2011) and (b) there is evidence that depression is correlated to anxiety in patients with heart disease (Frasure-Smith et al., 1995, Watkins et
al., 2013) and this comorbidity can affect between 21% to 26% of the patients (Frasure-Smith and Lesperance 2008, Doering et al., 2010). Both explanations support the validity of the A-CDS.

The mean age of the patients in our study was 54 years. Considering that the vast majority of these patients is eligible to enroll in cardiac rehabilitation, their age can be compared with the age of patients who are enrolled in cardiac rehabilitation worldwide. Data from national cardiac rehabilitation registries report an average enrollment age of 67 years in the UK (Doherty et al., 2018), 64 years in Germany (Bestehorn et al., 2011), and 66 years in Canada (Grace et al., 2015). Additionally, the percentage of male patients in our study was 85.4%, compared to 71% in the UK (Doherty et al., 2018), 71.7% in Germany (Bestehorn et al., 2011), and 71% in Canada (Grace et al., 2015). Since age and sex influence depressive symptoms, the above differences in age and sex might render the results of our study not generizable to those countries.

It should also be noted that half of the sample in our study were inpatients with acute coronary syndrome, which means that these patients were assessed for depressive symptoms shortly after they suffered a cardiovascular event. Generalizing the results of our study to out-patients with heart disease, who may have overcome the acute psychological impact of a cardiovascular event, should be done with caution.

**Limitations**

The sensitivity and specificity of the A-CDS were not objectively assessed because it was not possible to perform a clinical evaluation of the sample.

**Conclusions**

The A-CDS appears to be a valid and reliable instrument that measures depressive symptoms in Arab in-patients with heart disease. The spectrum of nationalities and the typicality of the diagnoses in the sample encourage the generalization of the findings to the vast majority of Arab patients with heart disease. Further studies to assess sensitivity and specificity of the A-CDS are needed.
Conflict of interest

All authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to thank Rasha Khalil, Nermin Ibrahim, and Reem Alharbi for their valuable assistance in recruiting participants and administering the scales. Without their help this study would not be possible.

This study was funded by the Medical Research Centre of Hamad Medical Corporation (grant number 13206/13). The funding source had no involvement in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.
References


Hare, D.L., Meyer, K., McBurney, H., 2000. Cross-cultural robustness of depressed mood in cardiac patients — the German language version of the cardiac depression scale. Heart Lung Circ. 9, A167.


Chapter 6

Does depression reduce cardiorespiratory fitness? A quasi-experimental study
A. Introduction

We have established that there is a negative correlation between depression symptom severity and CRF in otherwise healthy individuals (Papasavvas et al., 2016b) and patients with heart disease (Papasavvas et al., 2017) and we have identified sex and BMI as moderators of this relationship in patients with heart disease (Papasavvas 2018). The negative correlation between depression symptom severity and CRF has clinical and prognostic implications, while it may also provide another indication of a biological link between these two variables. The physician may consider that depressed individuals may tend to be unfit, and thus tend to have increased risk of all-cause mortality (Kodama et al., 2009, Lee et al., 2010). Conversely, the physician may consider that unfit individuals may tend to be depressed and this could influence assessment and treatment of the underlying disease. Depression screening might be considered, for example, or challenges in adherence to treatment, a common finding in depression (Grenard et al., 2011), might be anticipated. Furthermore, an inverse correlation between depression symptom severity and CRF encourages the prescription of exercise as a treatment modality for depression. Indeed, exercise is considered to be moderately more effective than a control intervention for reducing symptoms of depression (Cooney et al., 2013), although further research on the frequency, intensity, duration, and type of exercise is certainly needed.

However, despite the abovementioned implications of the correlation between depression symptom severity and CRF, it is still not clear if the correlation is a result of unidirectional causation, bidirectional causation or influence of other variable(s). Clarifying the nature of the correlation is important because it would shed light on the involved psychophysiological pathways and it may lead to more precise changes in clinical practice if causative pathways are uncovered. Specifically, if depression would reduce CRF without the mediation of other factors (such as physical activity) this would mean that depression has a direct physiological effect on CRF and therefore, depression would not only be considered a mental disease but also as a somatic disease. This could help alleviate the stigma associated to mental disease and encourage depressed individuals to seek help. Of the few studies that address the relationship between depression and CRF, only one has assessed
whether depression reduces CRF (Hollenberg et al., 2003). In this study, which included older female participants, the presence of depression for 2 – 4 years was associated with a reduction in CRF that was proportional to depression symptom severity. However, in order to assess causality between two variables one needs to control (or, at least, measure) the influence of other variables; and in this study, physical activity, which is a strong determinant of CRF (Laukkanen et al., 2009, Lakoski et al., 2011), was not controlled or measured.

Physical activity can be assessed using subjective and objective measures. Subjective measures include questionnaires in which physical activity during a period (e.g. one week) is recalled or logged in a diary. Objective measures include activity monitors, pedometers, and heart rate monitors. A recent review of reviews concluded that, although both types of measures have their limitations, objective measures of physical activity demonstrate less variability when they are compared to direct observation or the doubly labelled water technique (Ainslie et al., 2003) than subjective measures of physical activity (Dowd et al., 2018). Objective measures of physical activity have also been reported to be more valid and reliable than subjective measures in patients with heart disease (Alharbi et al., 2017). Consequently, an objective measure of physical activity (activity monitor) was selected in our quasi-experimental study that aimed to assess whether depression reduces CRF in patients with heart disease.

**B. Changes in the quasi-experimental study**

Having the Arabic version of the CDS ready (Papasavvas et al., 2016a), we prepared the research proposal for the study assessing whether depression reduces CRF in patients with heart disease. This was a multi-center proposal, including institutions and data collection in Qatar and in Greece. However, during the final steps of the research proposal submission we were informed by the medical research center of my facility in Qatar that, according to its policy, all foreign collaborating institutions should have been registered with Qatar Supreme Council of Health. The collaborating institute from Greece was the Medical School of University of
Athens and it was not registered with Qatar Supreme Council of Health. The registration process would have been lengthy and would have required several approvals and signatures, including the signature of the Dean of the University of Athens. As such, the collaborating research team in Athens, including an experienced professor with more than 200 publications, was not willing to engage in such a process. Consequently, the collaboration with the research team in Greece was not feasible and the research proposal was not submitted.

Without the collaboration with the research team in Greece the data collection was going to be conducted solely in Qatar and it would have taken too long to recruit the required sample of participants, which consisted of patients with heart disease who were recently diagnosed with depression – the justification of this inclusion criterion will follow. Therefore, we decided to broaden the inclusion criteria and conduct the study in healthy individuals who were recently diagnosed with depression.

The proposed study aimed to be the first to assess whether depression reduces CRF in men, while controlling or objectively measuring the influence of potential moderators and/or mediators. The study was also relevant to Qatar, as its results would potentially help to improve services and treatment for depression, which was one of the goals and objectives of Qatar National Research Strategy, as stated in its Health Pillar "HE.1.9: Develop and implement research programs to reduce the occurrence and improve services and outcomes of mental health conditions, brain injury and epilepsy" (Qatar National Research Fund, 2012).

C. Research design for the quasi-experimental study

The quasi-experimental study aiming to assess whether depression reduces CRF in men would adopt the following design:

C.1 Objectives

i. To compare CRF, expressed by maximal oxygen consumption (VO2peak), between baseline recent (less than one month ago) major depressive disorder (MDD) episode and remission or six months (whichever
comes first), adjusting for changes in depression symptom severity, BMI, smoking status, physical activity, total sleep time, medication, and lifetime MDD episodes, in depressed men and matched healthy controls.

ii. To assess the effect of depression on respiratory and hemodynamic variables, including oxygen consumption (VO\textsubscript{2}) at the ventilatory threshold, minute ventilation to carbon dioxide production (VE/VCO\textsubscript{2}) slope, partial pressure of end-tidal carbon dioxide (PETCO\textsubscript{2}), O\textsubscript{2} pulse at VO\textsubscript{2PEAK}, ratio between O\textsubscript{2} and workload (\Delta VO\textsubscript{2}/\Delta W), maximum decrease in pulse oximetry (SpO\textsubscript{2}), maximum heart rate, and heart rate recovery.

iii. To assess differences in the abovementioned respiratory and hemodynamic variables between depressed men and matched healthy controls.

The hypothesis would be that depression reduces CRF and the anticipated outcome would be a reduction in CRF over time that could be attributed to depression after adjusting for the abovementioned potentially confounding variables.

C.2 Study methodology

i. Type of study

The type of the study is non-therapeutic clinical research; it is a quasi-experimental study with matched control group. Patients with unipolar depression and healthy controls matched for age, gender, BMI, smoking status, and physical activity would be recruited. CRF would be compared between baseline recent (less than one month ago) MDD episode and remission or six months ( whichever comes first).

ii. Setting & Location

Patients would be recruited in Hamad Medical Corporation (HMC) Psychiatry Department. Data will be collected in the out-patient psychiatric clinics of the Psychiatry Department and the cardiopulmonary exercise test (CPET) room of the Heart Hospital.
iii. Sample size calculation

The total number of participants would be 86. There would be 43 male adults with unipolar depression and 43 healthy controls matched for age, gender, BMI, smoking status, and physical activity. Based on published data (Donath et al., 2010, Krogh et al., 2012, Valtonen et al., 2009), patients with depression are expected to have a mean of maximal oxygen consumption ($\text{VO}_{2\text{PEAK}}$) around 30 mL min$^{-1}$kg$^{-1}$ with a standard deviation around 7.5. The minimum clinically important reduction in CRF will be set to 10%, which in this case is 3 mL min$^{-1}$kg$^{-1}$, yielding an effect size of 0.4. We would include eight covariates (depression symptom severity, BMI, smoking status, metabolic equivalents of task (MET)s, times spent performing moderate to vigorous physical activity (MVPA), total sleep time, medication, and number of lifetime MDD episodes) that will be assumed to account for 10% of the total variance. Based on these assumptions, a two-sided type 1 error probability of 5%, and a power of 90%, the total sample size should be 68 participants (calculated using G*Power ver. 3.1.9.2). Dropouts in MDDs can vary from 26-36% (Machado et al., 2006). We would anticipate a dropout rate of 30% in the depressed participants and 10% in the healthy controls; therefore, a mean of 20% dropout rate would be anticipated, yielding a required total sample size of 86 participants.

iv. Interventions involved in the research

The following interventions would be performed in all participants:


b. Telephone administration of the short form of the International Physical Activity Questionnaire (IPAQ) (Ainsworth et al., 2000).

c. Administration of the Hamilton Depression Scale (HAM-D) (Hamilton 1960).
d. Administration of the ActiGraph activity monitor

(https://www.actigraphcorp.com/products/wgt3x-bt-monitor/).

e. Conduction of CPET.

v. Data to be collected

Screening: The DSM-5 would be administered to establish the diagnosis of unipolar depression. The short form of the IPAQ would be administered via telephone to assess the physical activity level (see C.3.i Inclusion criteria).

Baseline and remission or six months (whichever comes first):

a. Assessment of CRF by measuring VO\textsubscript{2\text{PEAK}} during CPET; this variable would be the primary outcome.

b. Assessment of other respiratory and hemodynamic variables, including VO\textsubscript{2} at the anaerobic threshold, VE/VCO\textsubscript{2} slope, PETCO\textsubscript{2}, O\textsubscript{2} pulse at VO\textsubscript{2\text{PEAK}}, ΔVO\textsubscript{2}/ΔW, maximum decrease in SpO\textsubscript{2}, maximum heart rate, and heart rate recovery; these variables would be the secondary outcomes.

c. Assessment of anthropometric and basic life signs, depression symptom severity, smoking status, METs per day, time spent performing MVPA, and total sleep time during last week, medication list, and number of lifetime MDD episodes; these variables would serve as covariates in the statistical analysis.

The second CPET would be performed at remission or six months (whichever comes first). In case of remission, CRF should be measured before it potentially improves due to remission. For this reason, depression symptom severity would be assessed at least once per month so that potential remission would be documented promptly and the CPET would be performed
right after remission. In this case, any potential improvement to CRF due to remission would be minimal.

C.3 Study population

i. Inclusion criteria

Male adult patients aged 18-65 years with recent (less than one month) first MDD episode or male adult patients with unipolar depression who have suffered a recent recurrent MDD episode after a period of at least three months of remission and matched healthy controls for age, gender, BMI, smoking status, and physical activity would be included. Age, gender, BMI, smoking status, and physical activity have been reported to be the strongest determinants of CRF (Sandvick et al., 1995, Laukkanen et al., 2009, Lakoski et al., 2011). The recent MDD episode would serve to minimize any already existing effect of depression on CRF. In case of recurrent MDD episode, the period of at least three months of remission would also serve to minimize any previous effects of depression on CRF. Participants would be able to understand, speak, and read English or Arabic. The short form of the IPAQ would be administered via telephone to the participants and participants whose physical activity was “low”, using to the categorical scoring of the scale, would be included. This criterion would serve to reduce changes in physical activity over time, since sedentary individuals seem to be less likely to become active than vice-versa (Chen and Millar 2001). Participants would also be asked if they changed the level of their physical activity during the last month and they would be included if they did not change it. This criterion would serve to minimize any detraining effects of a potential decrease of physical activity on CRF. Diagnosis of unipolar depression would be confirmed by means of a Structured Clinical Interview for DSM-5 axis disorders (American Psychiatric Association 2013) conducted by psychiatrist.
ii. Exclusion criteria

Patients with current drug or alcohol abuse, contraindications to physical exercise, current/previous psychotic or manic symptoms, and cardiovascular and pulmonary disease would be excluded. This piece of information would be retrieved from patient's file.

C.4 Study procedures

Participants would be required to pay two baseline visits at the Heart Hospital and then monthly visits until remission or six months (whichever comes first) at the HMC Psychiatry Department. At remission or six months (whichever comes first) participants would last be required to pay two visits at the Heart Hospital. Specifically, participants would undergo the following procedures:

i. Screening for participation: An exercise physiologist/physiotherapist will administer the IPAQ via telephone (telephone script can be seen in Appendix) and score it using the categorical scoring. This will be performed at the Heart Hospital.

ii. Baseline assessment: An exercise physiologist/physiotherapist will explain the use of the activity monitor to the participant and he will attach it on participant’s non-dominant wrist. The participant would be instructed to wear the accelerometer for seven days continuously, without taking it off. This procedure would be performed at the Heart Hospital. After this seven-day period the participant would come again to the Heart Hospital to return the accelerometer and undergo the following procedures:

a. Anthropometric and basic life signs assessment, smoking status recording, requisition of patient’s medication report and number of lifetime MDD episodes from patient's file. The collection of this data would be conducted by a nurse.

b. Administration of the HAM-D. A psychiatrist, who would be a member of the research team, would administer and score the HAM-D.
c. Conduction of the CPET. The CPET would be conducted by an exercise physiologist/physiotherapist and one nurse, under the supervision of a cardiologist. This staff would not have collected any other data and they would be blinded to participant group allocation.

iii. Monthly assessment of depressive symptoms: A psychiatrist, who would be a member of the research team, would administer and score the HAM-D once per month. This procedure would be performed at the HMC Psychiatry Department.

iv. Final assessment: The final assessment would be conducted at remission or six months (whichever comes first) at the Heart Hospital and it would be the same with the baseline assessment.

C.5 Informed consent

In patients with depression a psychiatrist, who would be a member of the research team, would request consent from the participant at the out-patient psychiatry clinic. In healthy controls the participant would be invited to the Heart Hospital, where a member of the research team would request their consent. Both patients and healthy controls would be informed of the study details, benefits and risks, and they would have any question answered. They would then be given a period of three days to decide on their participation. The informed consent would be signed by the participant and signed and dated by the researcher who provided it and it would be stored in the PI’s locked drawer that was dedicated to the study. A copy of the informed consent would be given to the participant and another copy would be uploaded to patient’s electronic file. The consent was also translated to Arabic by the Arab physicians of the research team.

C.6 Anticipated risks

Anticipated risks are related only to the CPET. The CPET is generally safe and is performed without any problems in the vast majority of cases. However, as any medical procedure, it may have some rare
complications, such as low blood pressure, arrhythmias and, in extremely rare cases [1 in 20,000 tests (Stuart and Ellestad 1980)], death. Current guidelines state that, in most cases, exercise tests can be performed by sufficiently trained and competent non-physician health professionals and supported by a skilled physician who is in close proximity (Myers et al., 2014). However, in order to eradicate the already minimum risk, we would perform the CPETs under the supervision of a cardiologist who would be present in the room.

C.7 Data collection

All the following data would be collected at baseline and remission or six months (whichever comes first):

i. \( \text{VO}_{2\text{PEAK}} \) using CPET.

ii. \( \text{VO}_2 \) at the ventilatory threshold, VE/VCO\(_2\) slope, PETCO\(_2\), \( \text{O}_2 \) pulse at \( \text{VO}_{2\text{PEAK}} \), \( \Delta \text{VO}_2/\Delta W \), maximum decrease in \( \text{SpO}_2 \), maximum heart rate, and heart rate recovery using CPET.

iii. Depressive symptoms using HAM-D; METs per day, time spent performing MVPA, and total sleep time using activity monitor; BMI from nurse’s assessment; smoking status from patient; medication and lifetime MDD episodes from patient’s file. Depressive symptoms would also be collected once per month using HAM-D.

Cardiopulmonary indices would be collected by the exercise physiologist/physiotherapist who would conduct the CPET. Activity monitor indices would be collected by the same exercise physiologist/physiotherapist. HAM-D score would be collected by a psychiatrist. BMI, smoking status, medication list, and lifetime MDD episodes would be collected by a nurse. All data would be collected at baseline and remission or six months (whichever comes first). HAM-D score would also be collected once per month. All data would be collected at the out-patient clinics of HMC Psychiatry Department and at the CPET room of the Heart Hospital.

C.8 Statistical analysis

A mixed-design ANCOVA with the following parameters would be conducted:
Within groups factor: Time, with two levels; baseline and remission or six months (whichever comes first).

Between groups factor: Health status, with two levels; depressed and healthy participants.

Primary outcome: VO$_{2\text{PEAK}}$.

Secondary outcomes: VO$_2$ at the ventilatory threshold, VE/VCO$_2$ slope, PETCO$_2$, O$_2$ pulse at VO$_{2\text{PEAK}}$, ΔVO$_2$/ΔW, maximum decrease in SpO$_2$, maximum heart rate, and heart rate recovery.

Covariates: HAM-D score, BMI, smoking status, METs per day, time spent performing MVPA, total sleep time, medications, and lifetime MDD episodes.

BMI, smoking status, METs per day, time spent performing MVPA, total sleep time, medications, and lifetime MDD episodes were not expected to change significantly; therefore, they might fit as covariates in the ANCOVA analysis. HAM-D score might also fit as covariate if it did not change significantly or if it significantly correlated (covariates) with CRF. If, however, HAM-D score changed significantly and/or did not correlate (covariate) with CRF then HAM-D’s score suitability to be included as a covariate in the ANCOVA might be challenged, in which case a multiple regression analysis would be performed to determine predictors of CRF. Last, if significant correlations between covariates and CRF were found mediation analysis (MacKinnon, Fairchild et al. 2007, Valeri and Vanderweele 2013) would be conducted to explore potential mediating pathways.

The statistical analysis would be carried out by the PI (Theodoros Papasavvas) under the supervision of Prof. Dominic Micklewright who is an experienced researcher with advanced statistical analysis skills. Findings would be considered statistically significant if p < 0.05. Adjustments for multiple testing would be applied. Data would be analyzed using SPSS version 19 (SPSS, Inc., Chicago, USA).
C.9 Ethical issues

Participants would be paid QR200 when they would have completed their participation; that is, when they would have undergone the second CPET. The CPET is a symptom-limited exercise test that requires effort until fatigue is reached. Participants, especially patients suffering from depression, may be deterred from undergoing this procedure a second time. This would lead to dropouts and may threaten study completion. Providing financial incentive would help reduce dropouts and ensure study completion.

D. Delays in approval and conduction of the quasi-experimental study

Unfortunately, due to several circumstances, including delays in the transition from paper to paperless electronic proposal submission system, the selection of experienced reviewers, and the ethics application, the proposal was not approved until more than twelve months had passed. Moreover, after the proposal had been approved there was an extensive delay in purchasing and receiving the requested equipment for the study (accelerometers, in order to measure physical activity), which lasted around six months. The reasons for this delay were the fact that the manufacturer of the accelerometers was not an eligible vendor for my hospital, which resulted in looking for a third company through which the accelerometers could be purchased; and the change in climate, caused by the political, financial, and territorial embargo experienced in Qatar (Al Jazeera News and Agencies 2018). The embargo imposed on Qatar resulted in delays in the release of funds and in receiving the equipment due to the political, financial, and territorial blockade. Finally, after receiving the equipment and having the study approved, it proved unexpectedly difficult to persuade eligible patients to participate in the study, which delayed the data collection for six more months. Owing to time limits of the PhD, this study was regrettably abandoned for PhD purposes.
References


Thesis Discussion and Conclusions

Depression symptom severity was negatively correlated to CRF in otherwise healthy individuals and this relationship appeared stronger in men than in women. This finding has prognostic and clinical implications. A primary healthcare clinician who is treating a patient with elevated depressive symptoms may consider that this patient might have reduced CRF and therefore increased risk of mortality, especially if the patient is a man; therefore, the clinician may consider assessing patient’s CRF and may prescribe exercise, which would improve patient’s CRF and may also reduce depressive symptom severity. Conversely, a clinician who is treating a patient with low CRF may consider that this patient might exhibit elevated depressive symptoms, especially if the patient is a man; therefore, the clinician may consider screening the patient for depression and might anticipate challenges related to elevated depressive symptoms, including poor adherence to treatment.

Depression symptom severity was also negatively correlated to CRF in patients with heart disease and this correlation was moderated by sex and BMI: Lower BMI enhanced the correlation, while higher BMI attenuated the correlation in male patients and rendered it statistically not significant in female patients. This finding also has prognostic and clinical implications. A clinician who is treating a patient with heart disease and elevated depressive symptoms may consider that this patient might have reduced CRF, especially if the patient is a man with low BMI. Impaired CRF is an indicator for poor prognosis in patients with heart disease and therefore the clinician may consider conducting an objective assessment of patient’s CRF (i.e. a symptom-limited exercise test) and should prescribe exercise to improve CRF, unless contraindicated. Conversely, a clinician who is treating a patient with heart disease and impaired CRF may consider that this patient might exhibit elevated depressive symptoms, especially if the patient is a man with low BMI. Depression is a risk factor for poor prognosis in patients with heart disease and therefore the clinician may consider a more thorough assessment of patient’s depression symptom severity and refer the patient to a specialist if
required. The clinician may also be prepared to anticipate depression-related challenges, including poor adherence to treatment and cardiac rehabilitation.

Both meta-analyses included studies in which depression symptom severity was measured using psychometric scales. Seven scales were included in the first meta-analysis and ten scales were included in the second meta-analysis. Each of these scales has different range of scores and, although all the scales have been validated, their sensitivity in detecting depressive symptoms is probably not identical, especially in patients with heart disease, in whom generic depression scales have significant limitations in assessing depression (Vieweg et al., 2011). Both range of scores and scale sensitivity may have influenced the correlation between depression symptom severity and CRF in otherwise healthy individuals and patients with heart disease in our meta-analyses and therefore, they may have contributed to the heterogeneity of the results. Respectively, CRF in these meta-analyses was measured using symptom-limited tests. Studies in which CRF was measured using treadmill and cycle ergometer CPETs were included in the first meta-analysis, and studies in which CRF was measured using treadmill and cycle ergometer CPETs, simple treadmill tests, watts-based cycle ergometer tests, and ISWTs were included in the second meta-analysis. Although all these tests are symptom-limited, they may produce different results, which may influence the correlation between depression symptom severity and CRF. For example, the VO$_{2peak}$ derived by a treadmill test is 10% - 20% higher than the VO$_{2peak}$ derived by a cycle ergometer test (Myers et al., 1991). However, we found no influence of the type of exercise test on the correlation between depression symptom severity and CRF in patients with heart disease in our meta-analysis.

We found that the correlation between depression symptom severity and CRF was weaker in women than in men in otherwise healthy individuals. A potential explanation of this difference might be biological. Although the reasons for the female predisposition towards depression are still unclear (Eid et al., 2019), there is consensus that depression prevalence in women is at least 1.5 times its prevalence in men (World Health
Organization 2017). A potential clustering of depression symptom severity scores towards high values in women might have weakened the correlation between depression and CRF, whereas a potential more even distribution of depression symptom severity scores in men might have empowered the correlation. However, although the prevalence of depression in women with heart disease is similarly higher than in men with heart disease (Shanmugasegaram et al., 2012), we did not find a difference in the correlation between depression symptom severity and CRF between men and women with heart disease in our meta-analysis. Perhaps the presence of heart disease influenced the correlation in both sexes, while it may be possible that other factors, which are yet to be determined, may have also influenced the correlation, both in otherwise healthy individuals and patients with heart disease. More research is needed in this area.

Our two meta-analyses have similarities and differences. Both meta-analyses used the same methodology to assess the correlation between symptom severity and CRF except for two inclusion criteria: the population (otherwise healthy individuals vs patients with heart disease) and the type of exercise test (CPET vs all symptom-limited exercise tests, including CPET, simple treadmill test, watts-based cycle ergometer test, and ISWT). The choice to include all symptom-limited exercise tests in the second meta-analysis was based on the fact that the CPET is still underutilized in patients with heart disease (Guazzi et al., 2018) and therefore, excluding studies with other symptom-limited exercise tests would have led to a loss of a significant volume of data. The inclusion of studies that used any symptom-limited exercise test to assess CRF in patients with heart disease resulted in more studies and participants being included in the second meta-analysis than the first (16 vs 59 studies, 4039 vs 25733 participants). Heterogeneity was also higher in the second meta-analysis (33% vs 64%) but, contrary to the first meta-analysis, in which we were able to identify sex as a moderator, we were not able to identify any cause for this significant heterogeneity. However, it was not possible to assess the within-study level moderating effect of any potential moderator in the meta-analyses because raw data were not available for every study. As such, we conducted subgroup analyses to assess the between-studies moderating effect of few potential moderators for which data were available. It should be noted,
though, that subgroup analyses in meta-analyses that employ random effects models, which was the case in our meta-analyses, can suffer from low power (Borenstein and Higgins 2013) and therefore, statistically non-significant results should be interpreted with caution. In a subsequent moderation analysis study I found that the correlation between depression symptom severity and CRF was moderated by sex and BMI in patients with heart disease (Papasavvas 2018); however, to my knowledge, this is the only moderation analysis study on the correlation between depression symptom severity and CRF and it is limited by the fact that the participants were recruited from 1993 to 2002 and therefore, its results should be generalized with caution. Consequently, there is need for further research in this area.

We were not able to assess whether the correlation between depression symptom severity and CRF was causative. In order to assess causality, the influence of potential mediators should be controlled or, at least, objectively measured. There are several potential mediators of a causal relationship between depression symptom severity and CRF, including physical activity, sex, age, smoking history, BMI, socioeconomic status, and anxiety, as they all influence either or both depression and CRF (see indicative references for physical activity (Laukkanen et al., 2009, Roshanaei-Moghaddam et al., 2009), sex (Laukkanen et al., 2009, Shanmugasegaram et al., 2012), age (Stordal et al., 2003, Laukkanen et al., 2009), smoking (Luger et al., 2014, Siddall et al., 2017), BMI (Laukkanen et al., 2009, Pereira-Miranda et al., 2017), socioeconomic status (Richardson et al., 2015, Ombrellaro et al., 2018), and anxiety (Pollack 2005, Loprinzi et al., 2017). Of these potential mediators, physical activity could perhaps be the most plausible, given the fact that it is strongly associated with depression and CRF (Laukkanen et al., 2009, Shanmugasegaram et al., 2012) and its mediation could be expected: Depressed individuals may not be motivated to engage in physical activity (Schuch et al., 2017), which would reduce their CRF.

In our quasi-experimental study that aimed to assess whether depression reduces CRF we planned to assess physical activity objectively using accelerometers. The decision to use an objective measure of physical activity instead of self-reported scales was based on the fact that objective measures of physical activity
exhibit lower levels of variability in their validity and reliability than subjective measures (Dowd et al., 2018). Moreover, subjective measures of physical activity may not correlate strongly with objective measures (Prince et al., 2008). In fact, self-reported physical activity can be over-reported (Hartley et al., 2015) and may even be twice than the objectively measured physical activity using accelerometers (Colley et al., 2018). Sedentary time (Celis-Morales et al., 2012) and change in physical activity levels (Limb et al., 2019), which are both relevant to our quasi-experimental study, have also been reported to be measured with less precision using subjective measures of physical activity than accelerometers.

If depression reduces CRF this causation could be mediated by physical activity, as mentioned above, or it could be the result of psychobiological processes that are currently unknown. Respectively, such a causative relationship might imply that reducing depressive symptoms could increase CRF. Similarly, such a causation could be the result of currently unknown psychobiological processes or it could be mediated by physical activity: When depressive symptoms subside patients might be more encouraged to engage in physical activity and therefore, they may increase their CRF. There is evidence that depressed individuals engage in lower levels of physical activity and higher levels of sedentary behavior than healthy individuals (Schuch et al., 2017). Would relieving depressive symptoms make these patients more physically active and less sedentary? It should be noted, though, that in our quasi-experimental study we aimed to “control” physical activity by recruiting participants who were sedentary, as they have been reported to have less chance of changing their physical activity levels and become physically active than vice-versa (Chen and Millar 2001). A causative relationship between depression and CRF with physical activity controlled would imply the existence of psychobiological processes that are currently unknown.

Cardiac rehabilitation has long been considered to reduce all-cause and cardiovascular mortality (Taylor et al., 2004, Heran et al., 2011). However, there have been recent advances in the surgical and medical management of ischemic heart disease that have reduced mortality, including the introduction of primary percutaneous coronary intervention (Huynh et al., 2009), statins (Naci et al., 2013), ACE inhibitors (ACE
Inhibitor Myocardial Infarction Collaborative Group (1998), and clopidogrel (Chen et al., 2005), and it is a fact that mortality from ischemic heart disease has progressively decreased from 2005 to 2015 (Nowbar et al., 2019). It may therefore be logical to argue that systematic reviews on the mortality benefits of cardiac rehabilitation that included data from older studies may be overestimating the current mortality benefits of cardiac rehabilitation. Indeed, the 2016 update of Heran’s et al. Cochrane systematic review (Heran et al., 2011) that included 16 new randomized controlled trials (RCTs) concluded that cardiac rehabilitation reduced cardiovascular mortality but not all-cause mortality (Anderson et al., 2016), while a recent review that included RCTs whose participants were recruited after 2000 concluded that cardiac rehabilitation had no effect on all-cause or cardiovascular mortality (Powell et al., 2018). Another recent meta-analysis (Takura et al., 2019) reported similar findings. However, this meta-analysis focused on cost-effectiveness of cardiac rehabilitation and included only three studies in its mortality analysis; therefore, its methodology was significantly limited. On the contrary, a meta-analysis that included studies that were published in 1995 or later concluded that cardiac rehabilitation was associated with reduced all-cause mortality (Rauch et al., 2016); however, this meta-analysis also included non-randomized studies and the only RCT that was included showed no mortality benefit. A recent network meta-analysis concluded that cardiac rehabilitation that included exercise training and/or psychosocial management reduced all-cause mortality (Kabboul et al., 2018); however, this meta-analysis included all published studies until 2017 and therefore, did not account for the advances in the management of heart disease. The number of prescribed exercise sessions was found to be positively related to all-cause mortality in a meta-analysis of RCTs and non-randomized studies (Santiago de Araujo Pio et al., 2017), while adherence to the prescribed exercise sessions was found to be positively related to cardiovascular mortality in a meta-analysis of RCTs (Abell et al., 2017); although these two meta-analyses included studies from inception and therefore, did not account for the advances in the management of heart disease, they assessed the effects of adherence to and dose of cardiac rehabilitation on mortality. None of the previous meta-analyses assessed these effects; therefore, the actual dose of exercise received by
patients in those meta-analyses was unknown, which may have had significant impact on the efficacy of the intervention and the respective findings of the meta-analyses.

In the light of the findings of the above meta-analyses, there seems to be a shift in the focus of cardiac rehabilitation stakeholders, including the British Association for Cardiovascular Prevention and Rehabilitation, regarding benefits of cardiac rehabilitation from mortality towards hospital admissions and psychosocial health, including quality of life, anxiety, and depression (British Association for Cardiovascular Prevention and Rehabilitation 2017). The European Association of Cardiovascular Prevention and Rehabilitation has published a position on the importance of assessment and management of psychosocial risk factors, including low socio-economic status, social isolation, stress, type-D personality, depression and anxiety, in patients entering cardiac rehabilitation (Pogosova et al., 2015), while the American Heart Association has published a statement elevating depression to a risk factor for poor prognosis in patients with acute coronary syndrome (Lichtman et al., 2014). Consequently, studying and providing enhanced evidence on the psychosocial aspects of cardiac rehabilitation is relevant, valuable, and should be encouraged. This thesis provided detailed information on the relationship between depression and CRF, which are perhaps the most significant psychosocial and physical health targets, respectively, in cardiac rehabilitation.

We also conducted the translation and validation of the CDS to Arabic, which was warranted for the continuity of the PhD, creating a valid and reliable instrument that was validated in a representative sample of Arab patients with heart disease; therefore, the A-CDS can be administered to the vast majority of this population. Additionally, the A-CDS demonstrated a closer to normal distribution of scores compared to the positively skewed distribution of scores of the A-HADS-D, which enables it to detect even mild depressive symptoms, which have prognostic and clinical significance in patients with heart disease.

The topic of this thesis is relevant to Qatar, a country in which cardiovascular disease is the second cause of death (after injuries) (Chaabna et al., 2018) and 83% of its population is engaged in low or no physical activity
The findings of the thesis have been used to improve cardiac rehabilitation services in the country. The translation and validation of the CDS to Arabic has produced a tool that we now use to screen all Arabic-speaking patients with heart disease who are enrolled in in-patient and out-patient cardiac rehabilitation in Qatar, as depression is an established risk factor for poor prognosis in these patients (Lichtman et al., 2014) and depression screening should be part of their treatment and rehabilitation plan (Lichtman et al., 2008). Patients with high CDS scores, especially if they are male with low BMI, are considered more likely to have low exercise capacity and they may undergo CPET instead of regular exercise testing, as part of their assessment prior to enrollment to cardiac rehabilitation, in order for their exercise capacity to be objectively assessed using the respiratory exchange ratio ≥1.1 as a test termination criterion (Balady et al., 2010). Respectively, patients who have low exercise capacity, especially if they are male with low BMI, are considered more likely to have elevated depressive symptoms and they may undergo a more thorough assessment of their psychosocial health.

This thesis has several limitations. In both our meta-analyses we were not able to assess the effects of potential moderators of the correlation between depression symptom severity and CRF — such as age, sex, BMI, PA and medication — at the within-study level, because raw data were not available for every study; instead, we conducted between-studies moderator analyses for the potential moderators with available data. We included studies that measured depression symptom severity and exercise capacity using different scales and exercise tests, respectively, which may have increased heterogeneity. Additionally, only English-language studies were included in the analyses and, although publication bias was not significant, several studies were excluded because it was impossible to acquire the relevant data. In the moderator analysis study the data were collected between 1993 and 2002, patients who were referred to cardiac rehabilitation but were not enrolled or were enrolled without initial assessment of fitness were excluded from the study, and there was no record on physical activity or dietary factors; therefore, generalizing the findings of the study on the current population of patients entering cardiac rehabilitation should be done with caution. The sensitivity and
specificity of the A-CDS were not objectively assessed, as we were not able to conduct a clinical evaluation as a criterion measure, while our sample in this study consisted of in-patients with a mean age of 54 years; therefore, the generalization of the findings to older out-patients, which is the typical population in cardiac rehabilitation settings in other parts of the world, should be done with caution. Finally, we were not able to assess whether the correlation between depression symptom severity and CRF is causal.

Looking forward, more research is warranted on the relationship between depression symptom severity and CRF. Potential moderators and/or mediators of this relationship should be explored in otherwise healthy individuals and patients with different types of heart disease, including CHD, heart failure, and congenital heart disease. Additionally, the effects of treating one variable on the other should be assessed. Finally, research should be conducted to determine whether the correlation between depression symptom severity and CRF is causal and to which direction(s).
References


ACE Inhibitor Myocardial Infarction Collaborative Group. 1998. Indications for ACE inhibitors in the early treatment of acute myocardial infarction: systematic overview of individual data from 100,000 patients in randomized trials, *Circulation*, 97(22), 2202-12.


Thesis References


Anthoine, E., Moret, L., Regnault, A., Sebille, V., Hardouin, J. B. 2014. Sample size used to validate a scale: a review of publications on newly-developed patient reported outcomes measures. *Health and Quality of Life Outcomes* 12, 2.


Shaw, L. J., Merz, C. N., Bittner, V., Kip, K., Johnson, B. D., Reis, S. E., Kelsey, S. F., Olson, M., Mankad, S., Sharaf, B. L., Rogers, W. J., Pohost, G. M., Sopko, G., Pepine, C. J. 2008. Importance of socioeconomic status as a predictor of cardiovascular outcome and costs of care in women with suspected myocardial ischemia. Results from the National Institutes of Health, National Heart, Lung and Blood Institute-
sponsored Women's Ischemia Syndrome Evaluation (WISE), J Womens Health (Larchmt), 17(7), 1081-92.
Smart, N. A., Murison, R. 2013. Rate of change in physical fitness and quality of life and depression following exercise training in patients with congestive heart failure. Congest Heart Fail, 19(1), 1-5.
guideline from the American Heart Association and American College of Cardiology Foundation. 

Circulation, 124, 2458-73.


Williams, M. D., Harris, R., Dayan, C. M., Evans, J., Gallacher, J., Ben-Shlomo, Y. 2009. Thyroid function and the natural history of depression: findings from the Caerphilly Prospective Study (CaPS) and a meta-analysis. Clin Endocrinol (Oxf), 70(3), 484-92.


Appendix

Supplementary Material Fig. S2.1

Funnel plot of effect size by precision. CC correlation coefficient.
<table>
<thead>
<tr>
<th>Reference</th>
<th>k</th>
<th>CC</th>
<th>Cl-</th>
<th>Cl+</th>
<th>z</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hollenberg et al. 2003 [45]</td>
<td>15</td>
<td>-0.17</td>
<td>-0.22</td>
<td>-0.11</td>
<td>-5.39</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Stewart et al. 2003 [53]</td>
<td>15</td>
<td>-0.15</td>
<td>-0.21</td>
<td>-0.09</td>
<td>-5.01</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Kiive et al. 2004 [47]</td>
<td>15</td>
<td>-0.15</td>
<td>-0.21</td>
<td>-0.09</td>
<td>-5.05</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>St Clair Gibson et al. 2006 [52]</td>
<td>15</td>
<td>-0.16</td>
<td>-0.21</td>
<td>-0.10</td>
<td>-5.23</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Hoffman et al. 2008 [44]</td>
<td>15</td>
<td>-0.15</td>
<td>-0.21</td>
<td>-0.10</td>
<td>-5.11</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Boettger et al. 2009 [40]</td>
<td>15</td>
<td>-0.19</td>
<td>-0.22</td>
<td>-0.15</td>
<td>-9.56</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Chu et al. 2009 [41]</td>
<td>15</td>
<td>-0.15</td>
<td>-0.20</td>
<td>-0.09</td>
<td>-5.21</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Krogh et al. 2009 [49]</td>
<td>15</td>
<td>-0.16</td>
<td>-0.22</td>
<td>-0.09</td>
<td>-4.92</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Valtonen et al. 2009 [54]</td>
<td>15</td>
<td>-0.13</td>
<td>-0.17</td>
<td>-0.08</td>
<td>-4.97</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Donath et al. 2010 [43]</td>
<td>15</td>
<td>-0.15</td>
<td>-0.21</td>
<td>-0.09</td>
<td>-5.17</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Rice et al. 2010 [27]</td>
<td>15</td>
<td>-0.15</td>
<td>-0.21</td>
<td>-0.09</td>
<td>-4.81</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Krogh et al. 2012 [50]</td>
<td>15</td>
<td>-0.16</td>
<td>-0.21</td>
<td>-0.09</td>
<td>-5.00</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Raso et al. 2012 [51]</td>
<td>15</td>
<td>-0.16</td>
<td>-0.21</td>
<td>-0.10</td>
<td>-5.34</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Dery et al. 2013 [42]</td>
<td>15</td>
<td>-0.16</td>
<td>-0.21</td>
<td>-0.10</td>
<td>-5.33</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Hunter et al. 2013 [46]</td>
<td>15</td>
<td>-0.16</td>
<td>-0.21</td>
<td>-0.10</td>
<td>-5.30</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Kreuzfeld et al. 2013 [48]</td>
<td>15</td>
<td>-0.16</td>
<td>-0.22</td>
<td>-0.11</td>
<td>-5.78</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Results of exclusion sensitivity analysis. k number of remaining studies, CC summary correlation coefficient, Cl 95% confidence interval.
Supplemental Digital Content 3.1

The 34 articles listed below also assessed the relationship between depressive symptoms and exercise capacity in patients with heart disease. Of these, 26 studies included patients with coronary artery disease and eight included patients with heart failure. These articles were not included in the analysis because it was not possible to obtain the necessary data from the publications or the authors.

References

11. Doe VH. *Effect of cardiac rehabilitation in depressed versus nondepressed individuals on fitness, depression, and perceived exertion.* Ann Arbor, Kent State University; 2009.


Protocol Title: Does Depression Reduce Cardiorespiratory Fitness? A Quasi-Experimental Study
Application No.: 16185/16
Principal Investigator: Theodoros Papasavvas

For patient with depression:
Hello Mr. [patient’s surname], my name is [investigator’s full name], I am exercise specialist/physiotherapist and I am calling you from the Heart Hospital regarding the research study that you have discussed with your psychiatrist, Dr. [psychiatrist’s name]. Dr. [psychiatrist’s name] informed us that you would like to know more about our study.

The purpose of our research study is to examine if depression reduces physical fitness. We will recruit patients with recent depression and we will measure their fitness using a treadmill exercise test twice; once at the beginning and one more time at remission or six months, whichever comes first. This process will not take more than an hour each time. Prior to this process we will give patients a watch to wear on their wrists for seven days in order to measure physical activity. Therefore, patients will need to visit the Heart Hospital four times in total. Additionally, patients will need to be seen monthly by their psychiatrist in the Psychiatry Department until remission or for six months, whichever comes first. These visits will normally not take more than half an hour each. Patients will be paid QR200 when they undergo the second exercise test.

Do you have any questions? Now that you have a basic understanding of the study, do you think you would be interested in participating?

If No: Thank you for your time.

If Yes: Very well. We now need to ensure that you are eligible to participate. I am going to ask you a few questions about your physical activity during the last seven days. This will take approximately 10 minutes of your time. The information that you will provide to me will be kept private and will only be seen by researchers in the Psychiatry Department and the Heart Hospital who are working in this study. If you do not agree to continue the phone call or if you would like to stop at any time, you are free to do so and this will not affect your care at the Psychiatry Department or the Heart Hospital.

Do I have permission to ask you these questions?

[Administration of the telephone version of the International Physical Activity Questionnaire (IPAQ)]

For control participant:
Hello Mr. [participant’s surname], my name is [investigator’s full name], I am exercise specialist/physiotherapist and I am calling you from the Heart Hospital regarding the research study that you would like to find out more about, as you mentioned in [your email]/[phone call] on [date of email/phone call].
The purpose of our research study is to examine if depression reduces physical fitness. We will recruit patients with recent depression and we will measure their fitness using a treadmill exercise test twice; once at the beginning and one more time at remission or six months, whichever comes first. This process will not take more than an hour each time. Prior to this process we will give patients a watch to wear on their wrists for seven days in order to measure physical activity. Therefore, patients will need to visit the Heart Hospital four times in total. We will also recruit healthy individuals, like you, who will match patients with depression in certain characteristics (e.g. age and weight) and they will undergo exactly the same procedures that I mentioned before. Therefore, they will also need to visit the Heart Hospital four times in total. All participants will be paid QR200 when they undergo the second exercise test.

Do you have any questions? Now that you have a basic understanding of the study, do you think you would be interested in participating?

If No: Thank you for your time.
If Yes: Very well. We now need to ensure that you are eligible to participate. I am going to ask you a few questions about your physical activity during the last seven days. This will take approximately 10 minutes of your time. The information that you will provide to me will be kept private and will only be seen by researchers in the Psychiatry Department and the Heart Hospital who are working in this study. If you do not agree to continue the phone call or if you would like to stop at any time, you are free to do so and this will not affect your potential future care at the Psychiatry Department or the Heart Hospital.

Do I have permission to ask you these questions?

[Administration of the telephone version of the International Physical Activity Questionnaire (IPAQ)]

Post-response communication for all participants:

If the participant is potentially eligible to participate:

Based on your replies, it appears that you may be eligible to participate in the study. Would you like us to arrange a meeting with a member of the research team to discuss the full details of your potential participation?

Would you like me to send you more details on the study to review them before the meeting?

Thank you for your time. If you need any additional information please do not hesitate to contact me. My name is [investigator's full name] and you can reach me at [phone number] and/or [email address].

If the participant is not eligible to participate:

Unfortunately, based on your replies, it appears that you are not eligible to participate in the study.

Thank you for your time. If you need any additional information please do not hesitate to contact me. My name is [investigator’s full name] and you can reach me at [phone number] and/or [email address].