Cost-Effectiveness of Vortioxetine Versus Escitalopram and Venlafaxine as First-Line Treatments of Moderate to Severe Major Depressive Disorder in Finland

Master thesis submitted as part of the European Master program in Health Economics and Management

Student: Eemil Karttunen
UiO student number: 621684
EUR student number: 428733
Thesis Supervisor: Eline Aas

University of Oslo, Faculty of Medicine, Department of Health Management and Health Economics

Erasmus University Rotterdam, Erasmus School of Health Policy and Management

Oslo,
June 28, 2020

Table of Contents

Declaration of Oath

"DECLARATION OF OATH

I hereby declare, under oath, that this master thesis has been my independent work and has not been aided with any prohibited means. I declare, to the best of my knowledge and belief, that all passages taken from published and unpublished sources or documents have been reproduced whether as original, slightly changed or in thought, have been mentioned as such at the corresponding places of the thesis, by citation, where the extent of the original quotes is indicated.

The paper has not been submitted for evaluation to another examination authority, nor has it been published in this form or another."

SIGNATURE	
June 17, 2020 EEMIL KARTTUNEN	24/6-20
date and signature of student	date and signature of supervisor

List of Abbreviations

AE:	Adverse event
CBA:	Cost-benefit analysis
CEA:	Cost-effectiveness analysis
CEAC:	Cost-effectiveness acceptability curve
CrI:	Credible interval
CUA:	Cost-utility analysis
EVPPI:	Expected value of partially perfect information
EVPI:	Expected value of perfect information
HAM-D:	Hamilton Rating Scale for Depression
HILA:	Lääkkeiden hintalautakunta [The Finnish Pharmaceutical Pricing Board]
HPA:	Hypothalamic-pituitary-adrenal (axis)
HRQoL:	Health-related quality of life
ICD:	International Statistical Classification of Diseases and Related Problems
ICER:	Incremental cost-effectiveness ratio
KELA:	Kansaneläkelaitos [The Finnish Social Insurance Institution]
MADRS:	Montgomery-Åsberg Depression Rating Scale
MC:	Monte Carlo (Simulation)
MDD:	Major Depressive Disorder
NMA:	Network meta-analysis
NMB:	Net monetary benefit
PSA:	Probabilistic Sensitivity Analysis
QALY:	Quality-adjusted life-year
RCT:	Randomized controlled trial
SAVI:	Sheffield Accelerated Value of Information (online tool)
SNRI:	Serotonin-norepinephrine reuptake inhibitor
SSRI:	Selective serotonin reuptake inhibitor
THL:	Terveyden ja hyvinvoinnin laitos [The Finnish Institute of Health and Welfare]
VOI:	Value of information (analysis)
WTP:	Willingness-to-pay

Abstract

Background: Major depressive disorder is a mental disorder characterized by apathy, sadness, and a lack of interest. The condition is commonly treated with psychotherapy and, especially in severe cases, antidepressants. Depression is increasingly recognized as a public health concern due to the economic burden it creates, as well as its role in suicide incidence. In Finland, treatment guidelines for first-line depression do not define a preferred antidepressant and cost-effectiveness literature is scarce. This study compares vortioxetine, a potentially underused and relatively novel treatment, with two of the most prescribed antidepressants on the Finnish market, escitalopram and venlafaxine, together with placebo.

Methods: This study conducts a cost-utility analysis to evaluate the cost-effectiveness of vortioxetine with the chosen comparators. The analysis includes a cohort Markov model, which was used to estimate costs, effects, and incremental cost-effectiveness ratios. In addition, scenario analyses and a probabilistic sensitivity analysis were conducted to assess the effect and magnitude of uncertainty, including the value of information. The parameters of the Markov model were informed by a structured literature review and a previous network meta-analysis consisting of relevant randomized controlled trials.

Results: The model predicted venlafaxine to be the least costly and most effective treatment alternative. In other words, it dominated the other treatments in the analysis. Vortioxetine was associated with the highest cost, although it was also associated with higher benefits than escitalopram. Only small differences were predicted between the active treatment alternatives, and all of them dominated placebo. The societal perspective indicated the majority of costs come as indirect costs through productivity losses. Finally, the sensitivity analysis indicated a substantial amount of uncertainty, as none of the alternatives had a probability of cost-effectiveness above 37%. Approximately €1343 per patient could be spent on additional research.

Conclusion: The study demonstrates that vortioxetine should be considered an equivalent treatment option to escitalopram and venlafaxine, and highlights the substantial effect of uncertainty in assessing treatments of depression. As none of the treatment alternatives emerge as definitively cost-effective over the others, the current treatment guidelines are considered justified in their lack of a decisive treatment choice.

1. Introduction

Major depressive disorder (MDD) is a mental disorder characterized by apathy, sadness, and a lack of interest in previously pleasurable activities. While sometimes challenging to separate MDD from "ordinary" mood changes, MDD is distinguished by considering persistence, as well as severity and possible other symptoms, such as decreased libido and changes in appetite (NCCMH, 2010). Furthermore, recovered MDD patients face, on average, approximately a 50% risk of relapsing after their first depressive episode (NCCMH, 2010). The condition is also associated with an increased risk of self-harm and a 20-fold risk of suicide (Osby et al, 2001). Despite the possibly dire consequences of MDD, and indeed its public health relevance, cost-effectiveness studies of depression treatments are scarce. This study aims to assess the cost-effectiveness of specific antidepressants in the treatment of adults with moderate to severe MDD in Finland.

The 12-month prevalence for MDD in adults in Finland is estimated at 7.4% (Markkula et al., 2015), and more than 280,000 people between ages 18 to 64 were prescribed antidepressants in 2019 (8.5% of population) (THL, 2020). A study by Saarni et al (2006) found that, after musculoskeletal disorders, psychiatric disorders were associated with the largest health-related quality of life losses in Finland on a population level. On an individual level, depressive disorders were found to have the third largest negative impact on quality of life, after Parkinson's disease and anxiety disorders. In addition, depression causes monetary costs to society: for example, according to the Finnish Centre for Pensions, mental disorders such as depression are the largest contributors towards early retirement on disability pension in Finland (2020).

The disorder is commonly treated with pharmacological treatment and/or psychotherapy (Huttunen, 2018). In Finland, selective serotonin reuptake inhibitors (SSRIs) such as escitalopram are most commonly prescribed, partly due to safety with regards to overdose as well as a favorable side effect profile as compared to previous generations of antidepressants, such as tricyclic antidepressants (Soini et al, 2017). Escitalopram and venlafaxine are the two most used antidepressants in the Finnish market (Fimea, 2019). Venlafaxine represents the class of serotonin-norepinephrine reuptake inhibitors (SNRIs), which have been found to potentially improve concentration difficulties, and are commonly prescribed for MDD patients in Finland (Fimea, 2019), as well as globally (McIntyre, 2017).

Vortioxetine received approval in 2013 (Ikäheimo, 2014). Besides being a potentially effective treatment, vortioxetine has been found to possibly improve cognitive functioning in MDD and has shown

effectiveness in relapse prevention (McIntyre, 2017). In addition, most patients experience relatively mild side effects. Most notably, rates of sexual disfunction are low (Baldwin et al, 2016). For these reasons, vortioxetine is recommended as a first-line treatment option in countries such as Canada (McIntyre, 2017). Currently, treatment guidelines in Finland do not specify which drug should be given priority in treatments, but it is left to the discretion of the practitioner. Vortioxetine is usually utilized as a third-line treatment after the failure of SSRIs.

Current cost-effectiveness literature in the Finnish context is scarce and, to the author's knowledge, there are no studies published comparing vortioxetine treatment with escitalopram in Finland. There is one cost-utility study comparing vortioxetine with venlafaxine after treatment switch in the treatment of MDD in Finland (Soini et al, 2017), and several studies in various contexts such as South Korea (Choi et al, 2016: Vortioxetine vs. venlafaxine XR) and Norway (Christensen & Munro, 2018: Vortioxetine vs. duloxetine) with different specifications regarding, for example, the measure of effect. This research adds to existing knowledge by making a comparison between vortioxetine, the current standards of first-line treatment (venlafaxine and escitalopram), and placebo, and is of specific significance for literature in the Finnish context.

After the introductory chapter, the thesis is structured as follows: a description of depression and the associated treatment methods is presented in Chapter 2, followed by a description of the theoretical framework in Chapter 3. Then, the research methods for this thesis will be elaborated on in Chapter 4, before presenting the results in Chapter 5. Finally, the significance of the results, the limitations of this study, as well as topics for further research will be discussed in Chapter 6 before arriving at a conclusion in Chapter 7.

2. Background

2.1. Definition of Major Depressive Disorder

Major Depressive Disorder, or simply, depression, is a systemic condition affecting both the mind and the body. Biologic theory has described depression as a neurochemical disturbance of monoamines, such as serotonin, noradrenaline, and dopamine (Karlsson, 2012). This theory has been supplemented by recent developments in neurobiological theory, with the description of a more complex model including, for example, the nerve growth factor and other elements in the central nervous system (Karlsson, 2012; Sotelo & Nemeroff, 2017).

The first depressive episode is triggered by stress in almost all cases (Karlsson, 2012). Stress triggers the hypothalamic-pituitary-adrenal (HPA) axis, which is responsible for adjusting a long-term stress response in the body through the so-called "stress hormone", cortisol. (Karlsson, 2012). Depression has been found to cause an overreaction of the HPA-axis, which releases an excessive amount of cortisol into the body. Increased cortisol levels may lead to shrinkage of the hippocampus, an area of the brain active in learning, memory, and controlling the HPA-axis (Karlsson, 2012). A decreased ability of the hippocampus to control the cortisol-excretion of the HPA-axis creates a negative feedback loop, causing an inability to deal with stress and an increased susceptibility to various comorbidities, such as cardiovascular diseases (Karlsson, 2012; Markkula et al, 2015; Sotelo & Nemeroff, 2017). Abnormalities in the functioning of the HPA-axis are a genetic predictor of the susceptibility to depression (Heiskanen et al, 2011). The severity of depression is nevertheless categorized by the patient's mental symptoms, rather than physiological properties.

2.2 Diagnostics

According to the Finnish Handbook of Psychiatric Classifications (THL, 2012), a depressed patient suffers from poor mood, as well as a lack of interest and energy for different activities. Other symptoms may include a decreased attention span, low self-esteem, feelings of guilt, insomnia, lack of appetite, and suicidal thoughts. A diagnosis requires that symptoms have been present for at least two weeks or are exceptional in their severity or sudden onset (THL, 2012). In addition, the patient should not have had previous manic or hypomanic episodes, and substance abuse should be ruled out as a cause of symptoms. The following table of diagnostic criteria is adapted from the Finnish guidelines for the treatment of depression and follows the International Classification of Disease (ICD) criteria (Duodecim, 2020).

Table 1: Symptomatic criteria for assessing depression severity

Criteria	Description
A. The depressive episode has lasted for at least two weeks	
B. At least two of the following symptoms are present	i. A depressed mood for most of the time
	ii. Loss of interest in previously pleasurable activities
	iii. Loss of energy
C. Some of the following symptoms are present: total number of symptoms (in B and C) add up to at least 4.	iv. Low self-esteem
	v. Unfounded feelings of guilt
	vi. Repeating thoughts of death or suicide or self-destructive behavior
	vii. Concentration difficulties, which may also present as decreased ability to make decisions
	viii. Psychomotor changes (excitement or lethargy)
	ix. Insomnia
	x. Changes in appetite, accompanied with weight loss or gain
Four to five symptoms are present in mild depression, six to seven in moderate, and eight to ten in severe, including all symptoms in B.	

2.2.1 Diagnostic Questionnaires

In addition to the symptomatic criteria presented in Table 1, several diagnostic questionnaires have been developed to help physicians and patients to accurately identify depression and its severity. Four are mentioned in the Finnish treatment and diagnostic guidelines (2016): the Montgomery-Åsberg depression rating scale (MADRS), the Beck Depression Inventory (BDI), the Hamilton Rating Scale for Depression (HAM-D), and the Patient Health Questionnaire-9 (PHQ-9). These questionnaires can be filled in by physicians or by the patients themselves and all contain similar questions to assess the severity of the symptoms listed in Table 1. The answers are used to create a score, which indicates the estimated severity of the patient's depression. If

the patient does not display a certain threshold score, or at least four of the symptoms presented on Table 1, (s)he will not be diagnosed with clinical depression (MDD).

2.3 Epidemiology

2.3.1 Incidence and Prevalence

Estimating the incidence of depression has been a challenge for research, and there is a lack of longitudinal studies (Ferrari et al, 2013). In addition, studies from different regions have used different methodology, making the results difficult to compare with each other. From a literature review, Ferrari et al (2013) estimate annual global incidence of depression at 3%. Nevertheless, the authors point out that the rather low incidence rate is not in line with prevalence rates established in different studies (including their own). This is because the average duration of a depressive episode is estimated at 30 weeks (i.e. less than a year), which suggests that the incidence rate should be higher than the prevalence rate (Ferrari et al, 2013). A cross-cultural study by Bromet et al (2011) found an average 12-month prevalence of 5.5% in high-income countries. Ferrari et al (2013) arrived at a similar figure of 5.4% as a global prevalence estimate (point or 12-month). Further, Markkula et al (2016) conclude based on a literature review, that the global 12-month prevalence of depression is approximately 5%, although they point to the significance of regional variance.

Markkula et al (2016) estimated, based on the Finnish Health 2011 study, that the 12-month prevalence of MDD in Finland in 2011 was 7.4%. Between 2000 and 2011, there was approximately a two-percentage point increase in MDD prevalence. Heiskanen et al (2011) estimate that on any given moment (i.e. point prevalence), 5-6% percent of the adult population in Finland suffer from depression. Further, 20% of the population is estimated to suffer from clinical depression at some point in their life (lifetime prevalence).

2.3.2 Prognosis

A Finnish study by Riihimäki et al (2014) found that 70% of their sample of MDD patients in primary care reached full remission during a 5-year follow-up with a median time of 20 months. A third of the sample had at least one recurrence. Heiskanen et al (2011) estimate that the average time to remission is between five and six months, and approximately 10% of patients develop into chronic cases. Heiskanen et al (2011) list several

factors that affect the length of recovery: current life circumstances; strong (weak) social network; personality traits; substance abuse; and, most importantly, severity of the depression.

2.4 Risk Factors

There are several risk factors associated with depression. Most depression cases are triggered by negative and stressful events in a person's life (Heiskanen et al, 2011; Karlsson, 2012). Examples include divorce, being laid off work, and the burden of taking care of a dependent family member. In addition, factors such as other somatic diseases, medical dependencies, addictions, or other mental disorders may contribute towards creating stress and act as additional risk factors. Young age is associated with a higher prevalence of depression (Markkula, 2016), although the relationship may vary across contexts (Bromet et al, 2011). Finally, there are biological factors, such as genetics and hormonal changes, which increase the susceptibility for depression (Heiskanen et al, 2011). Additionally, 10-15% of mothers experience a depressive episode shortly after giving birth. This is known as postpartum depression. Symptoms may develop during pregnancy, although usual onset of postpartum depression is within 3 months of giving birth (Heiskanen et al, 2011). This may also partly explain why women are approximately twice as likely to develop depression than men, although the reasons for this gender difference are debated (Kuehner, 2017).

2.5 Treatment

2.5.1 Finnish Health Care

The Ministry of Social Affairs and Health is responsible for organizing Finnish health care. Finnish citizens are covered by National Health Insurance through the Social Insurance Institution (Kansaneläkelaitos, KELA). Care is provided by municipalities on a primary level, whereas secondary care is provided in twenty hospital districts, and specialist care and research by five university hospital regions (STM, 2013). The Finnish system is publicly funded, and insurance payments are collected as part of income tax (STM, 2013).

Out-of-pocket payments are capped at specified levels for each health service provided. These payment levels are checked and adjusted every two years. For 2020-21, the out-of-pocket payment for a visit at a general practitioner is €20.60, which can be charged a maximum of three times within a calendar year (STM, 2019). Medical expenses are divided to three reimbursement categories: basic reimbursement (40%),

lower special reimbursement (65%), and higher special reimbursement (100%). These categories are determined by the Pharmaceuticals Pricing Board (HILA). Each patient is liable for a deductible payment of €50 each calendar year, before receiving a reimbursement for subsequent medical purchases, according to the relevant reimbursement category. There is an out-of-pocket expenditure cap on total medical expenses, which is set at €577.66 for 2020 (STM, 2020).

2.5.1.1 Treatment of Depression in Finland

Most MDD patients can be treated in primary care (Duodecim, 2020). Treatment of depression depends on the symptoms the patient displays and their severity, as well as the patient's own preferences. Common treatments include psychotherapy together or without antidepressants and, in severe cases, electroconvulsive therapy. Institutionalization may be considered if there is a high risk of suicidal behaviour (Heiskanen et al, 2011). Treatment in Finland is divided into three phases: acute care, follow-up care, and secondary prevention. The aim is to completely cure the patient in the acute care phase, and then prevent relapses in the short- and long-term by the follow-up and preventive phases, respectively. In practice, this means that remission is achieved during the acute care phase, which is defined as a patient no longer displaying the symptom criteria. After achieving remission, patients enter the follow-up care process for a period of at least 6 months. If the patient does not relapse during the follow-up period, (s)he is considered cured. In treatment-resistant, prolonged, cases, the period after remission is longer in order to ensure the recovery of the patient. This maintenance period lasts for at least 12 months, but can be substantially longer (Duodecim, 2020). Figure 1 illustrates below:

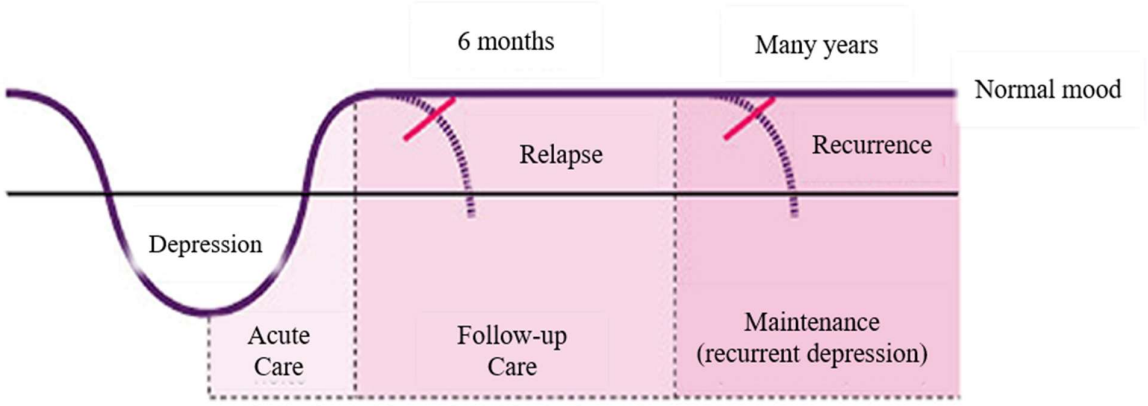


Figure 1: The three phases of treating depression. Adapted from Duodecim (2020)

2.5.2 Psychotherapy

Despite being a non-biologic treatment method, psychotherapy has been found to create changes in the neurobiological functioning of the patient's brain (Heiskanen et al, 2011). The key occurrence is an increase of serotonin receptor levels, which is also one of the main mechanisms of antidepressants. The aim of psychotherapy is to relieve stressful circumstances (e.g. relationship issues) and strengthen the capabilities of the patient to deal with hardships (Heiskanen et al, 2011). Psychotherapy can be executed in a short- or a long-form: short-form therapy usually consists of approximately twenty weekly meetings, whereas long-form therapy can take years. In total, there are hundreds of different types of psychotherapy, ranging from cognitive therapy to musical or interpersonal psychotherapy (Heiskanen et al, 2011). The details of these different forms are described elsewhere, but they all have the same goal, which is to help the patient cope with symptoms of depression and eventually achieve a symptom-free state.

2.5.3 Antidepressants

Antidepressants are a central part of depression treatments, especially in moderate and severe cases. Studies have shown that they are an effective treatment, however, challenges include poor patient adherence and identification of a suitable drug. Currently, physicians are unable to differentiate patients based on symptoms, biomarkers, or other factors (Heiskanen et al, 2011). Therefore, finding the optimal treatment regime is largely experimental. On average, more than 60% of patients show a response to treatment within 4-6 weeks, and half of them are relieved from all symptoms (Heiskanen et al, 2011). The use of antidepressants has increased over recent years (Fimea, 2019). However, this is also partly due to some drugs having multiple indications, thus being used for other conditions as well, such as anxiety or obsessive-compulsive disorder.

A network meta-analysis by Cipriani et al (2018) studied all antidepressants approved in the USA, Europe, and Japan before 2016. The study considers 522 double-blind studies including 116 477 randomly assigned patients and is arguably the most comprehensive comparative study on antidepressant efficacy to date. The analysis is referred to in the Finnish guidelines (2020) for the treatment of depression as "good quality" having "good transferability to Finnish context". The analysis by Cipriani et al (2018) found that all 21 antidepressants included were more effective than placebo, although with "modest" effect sizes.

2.5.3.1 Selective Serotonin Reuptake Inhibitors (SSRI)

More than 55% of the antidepressants consumed in Finland in 2018 were selective serotonin reuptake inhibitors (SSRIs) (measured in defined daily doses; THL, 2020). Their popularity is explained by a relatively favorable side effect profile, as well as their effectiveness in possible anxiety disorders, that approximately half of depression patients also suffer from (Heiskanen et al, 2011). Common side effects include nausea, diarrhea, and, for some drug types, sexual dysfunction. The most prescribed SSRI-antidepressant in Finland is escitalopram (brand names include Cipralex®), which is covered by basic reimbursement (40%) for MDD patients. Escitalopram may also be prescribed for psychotic disorders, for which it falls under the upper special reimbursement category (100%) (KELA, 2020). Other examples include fluoxetine (Seronil®), sertraline (Zoloft®), and citalopram (Cipramil®).

2.5.3.2 Serotonin-norepinephrine Reuptake Inhibitors (SNRI)

SNRIs may have better effectiveness as compared with SSRI due to their dual mechanism of action as both serotonin and norepinephrine reuptake inhibitors. The most commonly used drug in this class is venlafaxine (Efexor®), which is also the longest-standing SNRI on the Finnish market (Heiskanen et al, 2011) and has the same reimbursement status as escitalopram (KELA, 2020). Although possibly more effective than SSRIs, patients using venlafaxine may experience more severe adverse events, as well as eventual withdrawal symptoms. Additional examples of SNRIs include duloxetine (Cymbalta®) and mirtazapine (Remeron®).

2.5.3.3 Vortioxetine

Whereas SSRIs and SNRIs have single and dual mechanisms of action, respectively, vortioxetine is a multi-modal antidepressant. Its main targets are the serotonin receptors, or 5-HT receptors: vortioxetine functions as a 5-HT3A and 5-HT7 receptor antagonist, 5-HT1B receptor partial agonist, 5-HT1A receptor agonist, and inhibitor of the serotonin transporter (Citrome, 2014). Unlike escitalopram and venlafaxine, it gained approval relatively recently in 2013, and is not as established as a standard of care. It is patent protected and sold in Finland by Lundbeck under the trade name Brintellix®. Vortioxetine falls under basic reimbursement and does not have special status for any indication in Finland (KELA, 2020).

2.5.3.4 Other Antidepressants

Other pharmaceutical treatment methods include tricyclic antidepressants, which are the oldest antidepressants still in clinical use (Heiskanen et al, 2011). They are relatively affordable and display similar effectiveness to newer-generation treatments. However, patients can suffer from severe adverse effects, and overdose can be fatal. Examples include amitriptyline (Triptyl®), doxepin (Doxal®), and trimipramine (Surmontil®). In addition to tricyclic depressants, there are several other antidepressants that cannot be classified in the aforementioned categories. For example, agomelatine is an antidepressant that, in addition to serotonin receptors, targets melatonin (also known as the "sleep hormone") receptors (Heiskanen et al, 2011). Drugs such as agomelatine are, however, not the focus of this study.

3. Theoretical Framework

This section will describe the theoretical framework used for data collection and analysis in this thesis. In order to consider all relevant evidence, economic analyses need a way of synthesizing data. Therefore, systematic literature reviews and consequent meta-analyses are often conducted to support the analysis, as these methods can consolidate large amounts of data into statistically robust parameters to be used in an economic model. It may also be the case that treatment arms considered in a health economic evaluation have not been directly compared with each other in clinical trials, in which case a network meta-analysis might be needed to perform indirect comparisons. In addition, economic models themselves can be different in various ways with regards to, for example, measuring costs and effects, and statistical methods (e.g. extrapolation of data). Therefore, the concepts of meta-analysis and economic evaluation are described in this section.

3.1 Meta-Analysis

A meta-analysis is a way to synthesize data from multiple sources in order to incorporate all relevant evidence in the analysis. It is a method to combine data gathered from a systematic review and to create robustness to parameters in a decision analytic model, such as effectiveness estimates (Drummond et al, 2015). Whereas the model creates a framework to account for all relevant parameters in the decision such as cost, effectiveness, and resource use, a meta-analysis is a method to account for all relevant information in estimating any of those respective parameters. Despite the ability to improve precision, meta-analyses can also create misleading

results when biases and heterogeneity of the individual studies are not considered (Deeks et al, 2019). It is thus important to recognize that a meta-analysis will inevitably be of poor quality, if it refers to poor quality studies.

The first step of a meta-analysis is to collect (and calculate) the relevant summary statistic that can describe the effect of interest in the same way across studies (Deeks et al, 2019). Common examples include odds and risk ratios. After the necessary information is collected, a weighted average of the intervention effects from individual studies is calculated to arrive at a combined intervention effect estimate. The combination estimate may include an assumption about the true estimate of the individual studies. This is the distinction between fixed- and random-effect meta-analyses (Deeks et al, 2019).

A fixed-effects analysis assumes that the evidence relates to a common (true) effect. Variation between the observed effect in studies is therefore random, and the studies estimate the same underlying effect (Drummond et al, 2015). This assumption is challenged by heterogeneity between studies. Nevertheless, the goal in a fixed-effect analysis is to estimate the population effect. A random-effects approach, on the other hand, does not assume a common (true) effect, but rather that the effects estimated by the evidence share a common probability distribution. It is then assumed that the average of the estimates from different studies will be a reasonable estimate of the true effect. The estimated effects resulting from both approaches are often similar, but random-effect meta-analysis usually generates greater uncertainty than a fixed-effect approach, especially when that uncertainty is a reflection of heterogeneity (Drummond et al, 2015).

3.2 Network Meta-Analysis

A network meta-analysis (NMA) is a meta-analysis allowing for the comparison of multiple treatments through direct or indirect comparisons from randomized controlled trials (RCTs) (Hoaglin et al, 2011; Li et al, 2011). Whereas direct comparisons are based on clinical trials which compare the treatments of interest with each other, indirect comparisons are based on a common comparator across trials (Li et al, 2011). The resulting networks of evidence must not break randomization, meaning the comparisons should be based on one or more RCTs (Drummond et al, 2015). In addition, any bias resulting from the study identification method should be recognized and, if possible, eliminated. NMA may be conducted in order to consider all relevant evidence, answer questions without direct evidence, increase power and improve precision of estimates, or rank treatments (Hoaglin et al, 2011; Mavridis et al, 2015).

An indirect comparison between treatments X and Y can be made if they have each been compared with a third treatment Z. Comparing the absolute effects directly from each study (XZ and YZ) would break the randomization assumption. In order to retain randomization, the effect between XY (d_{XY}) can be estimated with the following equation (Drummond et al, 2015):

$$d_{XY} = d_{XZ} - d_{YZ} \quad (1)$$

An important assumption in NMA is transitivity, meaning that the distribution of the effect modifiers is expected to be the same across treatments. If the trial characteristics do not modify the effect of the interventions (e.g. age of participants), the transitivity assumption holds (Mavridis et al, 2015).

3.2.1 Frequentist and Bayesian Frameworks

NMAs are typically categorized into two frameworks: frequentist or Bayesian (Hoaglin et al, 2011). Most traditional statistical comparison methods follow frequentist methods, which provide point estimates and confidence intervals. Analyses involving indirect comparisons and thus more complicated models often employ Bayesian methods, which are computationally more intensive (Hoaglin et al, 2011). The Bayesian approach combines likelihood data with a prior probability distribution in order to estimate a posterior probability distribution of the parameters (Hoaglin et al, 2011). The prior distributions allow accounting for different sources of uncertainty, while the posterior probabilities allow for predictions, such as which treatment option is most likely to be most effective (Hoaglin et al, 2011). Results from a Bayesian analysis are presented with "credible intervals" (CrI), which provide an indication of uncertainty.

3.3 Economic Evaluation

Economic evaluation is a comparative analysis measuring and weighing the costs and consequences of two or more courses of action at a given point in time (Drummond et al, 2015). In the field of health care, these analyses are done for the purposes of informing decisions on the organization, funding, and execution of health care services in a given jurisdiction. It is important to note the perspective taken by the analysis, which differs considerably between, for example, multi-payer systems in the United States, and single-payer systems in Scandinavia (Drummond et al, 2015). In all cases, however, resources are scarce, and decisions on their

allocation needs to be informed by a systematic analysis. Economic evaluation allows for the incorporation of externalities and opportunity costs in (health care) decisions.

3.3.1 Types of Economic Evaluation

There are different types of economic evaluation. As listed by Drummond et al (2015), the main categories are cost-effectiveness analyses (CEA), cost-utility analyses (CUA), and cost-benefit analyses (CBA). They all measure costs in monetary terms, which implies their main differences are in the measuring of effects. In CEAs, effects are measured in natural units, such as life-years gained, or points reduced on a depression-rating scale. This specificity comes with limitations: a lack of comparability. From a CEA it may be difficult to assess opportunity costs when other programs falling under the same budget cannot be measured in the same measure of effect. A CUA, on the other hand, measures effects as healthy life-years, usually as quality-adjusted life-years (QALY), which allows for comparability of results across the health care sector. Finally, CBAs measure both costs and effects as monetary units (Drummond et al, 2015). Health outcomes can be translated to monetary terms through techniques assessing society's willingness to pay or productivity gained through disability days avoided, for example.

3.3.2 Quality-Adjusted Life-Year

To understand cost-utility analyses, the concept of quality-adjusted life-years (QALYs) must be explained. QALYs are a generic health measure, which measure the years lived in good health. Therefore, the measure can simultaneously account for gains in both longevity and quality of life. It is computed by multiplying the years lived in a given health state with a utility value representing the health-related quality of life (HRQoL):

$$QALY = HRQoL\ x\ years\ in\ health\ state \tag{2}$$

HRQoL is a utility value representing the quality of life in a specific health state. This utility value typically ranges from 0 (death) to 1 (perfect health) and describes the disease burden associated with a certain health state. These utility weights are determined through studies and questionnaires where respondents elicit a utility value that they associate with the health state in question (Drummond et al, 2015). There are many methods in which to conduct such a study, and the QALY-measure is not without its critics. However, QALYs remain

standard practice and the discussion about the measure's advantages and disadvantages can be found elsewhere.

3.3.3 Incremental Cost-Effectiveness Ratio

The incremental cost-effectiveness ratio (ICER) is a commonly reported outcome of a CUA. The ICER represents the incremental costs per an increment in health gained when comparing one treatment with another.

$$ICER = \frac{Cost\ of\ intervention - Cost\ of\ comparator}{Effect\ of\ intervention - Effect\ of\ comparator} \quad (3)$$

In a CUA, the effects will be measured in QALYs, although other measures, such as life-years or DALYs (disability-adjusted life years) can also be used to achieve an ICER. The ICER is often compared with a cost-effectiveness threshold, which represents the willingness-to-pay (WTP) for an incremental gain in health benefit. In the UK, for example, the range between £20 000 and £30 000 per QALY gained has been established as an actionable threshold. Nevertheless, exceptions are made, which has underlined the difficulty of determining a meaningful threshold value (Cleemput, 2011). This is also why many countries have not identified an explicit WTP-threshold, including Finland and other Nordic countries (Cleemput, 2011).

3.3.4 Net Monetary Benefit

Net monetary benefits (NMB) are a way to incorporate the WTP-threshold in the cost-effectiveness measure. The measure uses the threshold to present the difference in effects as a monetary value and, unlike the ICER, is a linear expression, which may be desirable in certain statistical analyses (Drummond et al, 2015). Calculating incremental net monetary benefits (INMB) for each treatment option allows for the construction of cost-effectiveness acceptability curves (see next sections).

$$NMB_x = Threshold \times Effect_x - Cost_x \quad (4)$$

$$INMB_{xy} = NMB_x - NMB_y \quad (5)$$

3.4 Uncertainty

The ICER and NMB measures are useful measures of cost-effectiveness, however, as such they fail to capture uncertainty. Healthcare is characterized by uncertainty of treatment outcomes, and thus costs and effects.

Therefore, cost-effectiveness measures are always accompanied with a level of uncertainty. The necessity of addressing this uncertainty in public investments such as health has been challenged by some authors, such as Arrow and Lind (1970), who argue that it is sufficient to address decisions based on expected value. They argue that, when the population is large, the expected value closely approximates the willingness to pay for decision alternatives with uncertain returns. Their theory is known as the so-called Arrow-Lind principle.

Proponents of uncertainty analysis, such as Briggs et al (2011), have countered the Arrow-Lind principle with three main arguments; (1) models are often nonlinear with multiplicative parameters; (2) there are opportunity costs associated with decision making, and; (3) it may be difficult and costly to reverse decisions. In addition, uncertainty analysis provides the tools for assessing value of information (see sections below), which can be a valuable tool in guiding future research. Uncertainty analysis is also recommended in the Finnish guidelines for conducting economic evaluations (Fimea, 2012).

3.4.1 Sensitivity Analysis

Sensitivity analyses are a way to assess uncertainty in a model, which can be categorized into two main types: parameter and structural uncertainty (Drummond et al, 2015). Parameter uncertainty refers to uncertainty in the inputs of the parameters in the model (e.g. cost and effectiveness parameters). Structural uncertainty relates to the assumptions made when building the model, such as the choice of time horizon or cycle length (Drummond et al, 2015).

Sensitivity analyses can be deterministic or probabilistic. A deterministic analysis (e.g. pessimistic scenario analysis) is generally not sufficient to address uncertainty, as it represents events that are extreme and highly unlikely. A deterministic analysis serves better as an indicator of the range in which results can vary (Fimea, 2012). A probabilistic sensitivity analysis (PSA) samples each uncertain parameter with an appropriate distribution and records the result with each set of parameters. This is repeated multiple times (e.g. 10 000) to achieve a likely range and distribution of outcomes (Drummond et al, 2015). A probabilistic analysis can be presented in the form of a cost-effectiveness acceptability curve or as a scatterplot on the cost-effectiveness plane. A PSA is the preferred form of uncertainty analysis in the Finnish guidelines for economic evaluations (Fimea, 2012).

3.4.2 Cost-Effectiveness Plane

A useful way to present cost-effectiveness data is as a scatterplot on the cost-effectiveness (CE) plane. The x-axis on a CE-plane represents the incremental effect of the intervention (i.e. the denominator of an ICER) and the y-axis represents incremental costs (i.e. the numerator of an ICER). A straight line is drawn through the origin, which represents the WTP-threshold. The simulated ICERs from a PSA are plotted on the plane, and all the estimates falling below the threshold-line are considered cost-effective with regards to the particular threshold (Drummond et al, 2015). Figure 2 illustrates:

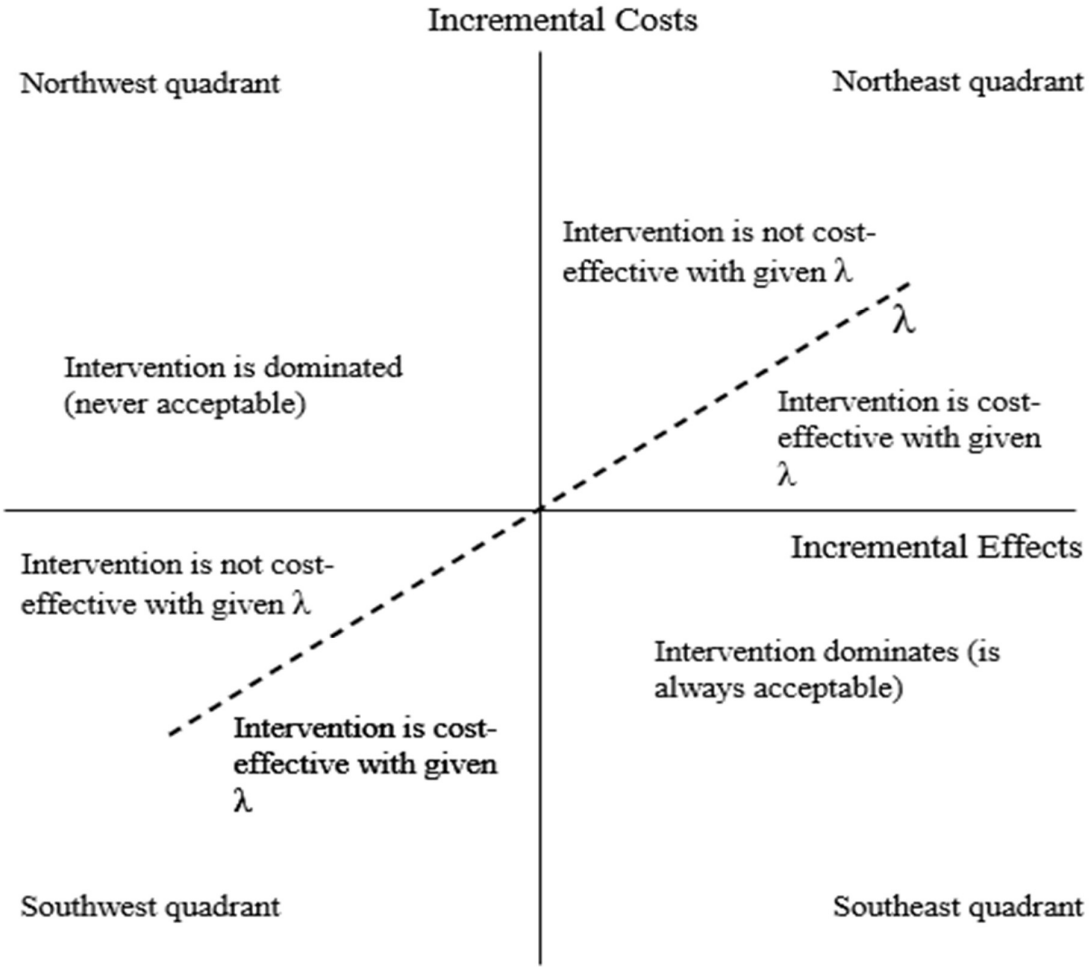


Figure 2: Cost-Effectiveness Plane. λ represents the WTP-threshold.

The CE-scatterplot gives an indication of the uncertainty associated with the ICER (spread), as well as whether that uncertainty is driven by costs or effects.

3.4.3 Cost-Effectiveness Acceptability Curve

A cost-effectiveness acceptability curve (CEAC) plots the probability of a treatment being cost-effective as compared to the other treatment alternatives under consideration (according to a PSA) against a range of cost-effectiveness thresholds. This enables a straightforward inspection of the effect of uncertainty on the probability of making a treatment recommendation that is (not) cost-effective. The probability represents the proportion of simulations where a given treatment has the highest net benefit in relation to the comparators (Drummond et al, 2015). A CEAC can provide an easy-to-interpret visualization of cost-effectiveness. However, in some cases the treatment with the highest probability to be cost-effective may not be the treatment with the highest expected net benefit (Drummond et al, 2015). This can occur when the mean value of the expected net benefits at different thresholds is higher than the respective median (Drummond et al, 2015).

3.4.4 Value of Information -analysis

A value of information (VOI) analysis can be used to determine which parameters to target with additional studies and whether those studies would be worthwhile. It is thus a way to estimate the return on investment of additional research. Expected value of perfect information (EVPI) is a common VOI-measure, which is defined as the difference between the expected net benefit with perfect information (i.e. no uncertainty) and the expected net benefit with current information regarding the uncertain parameters (θ). The output of a PSA can be used to calculate the EVPI with the following formula (Drummond et al, 2015):

$$EVPI = E_\theta max_i NB(i,\theta) - max_i E_\theta NB(i,\theta) \quad (6)$$

The EVPI calculation results in an upper bound on the expected benefits of attaining additional information regarding an individual patient. It can thus be used as a necessary indication of whether further research would be worthwhile. Accordingly, to achieve an upper bound for the whole population, the EVPI is multiplied by the population affected by the decision during a relevant time horizon. This is known as the population EVPI (Drummond et al, 2015).

To indicate where future research should be directed, one should use the same calculation principles to achieve the expected value of partially perfect information (EVPPI). The EVPPI is calculated as the difference between the expected net benefit with perfect information regarding a certain parameter and the

expected net benefit with current information about that same parameter (Drummond et al, 2015). An example of an application of EVPPI would be to compare the EVPPI values of efficacy parameters and safety parameters: a higher EVPPI suggests a greater potential of reducing decision uncertainty through additional research in that specific area.

EVPPI calculations can be conducted with similar methods as a standard PSA, with repeated sampling according to a probability distribution of uncertain variables. However, EVPPI calculations can be computationally very intensive due to the requirement of both an inner and outer loop of sampling (Drummond et al, 2015; Tuffaha et al, 2016). Therefore, Monte Carlo simulations commonly used for PSAs may be prohibitively slow and inefficient. An alternative to Monte Carlo methods is to use the result of a PSA in a nonparametric regression approach in R software (R Foundation, Vienna, Austria), as described by Strong et al (2014). This approach can significantly relieve the computational burden. The code is made available by Strong et al through the Sheffield Accelerated Value of Information tool online (2014).

3.5 Model Validity

Validation is a term used when assessing the ability of a model to reproduce reality (Eddy et al, 2012). It is a way to achieve trust and confidence in a health economic model and is thus a key feature of good modeling practices. While sensitivity analyses may account for variation in inputs (uncertainty), they do not evaluate how accurately the model portrays reality. Therefore, sensitivity analyses and validation can be considered as complements to each other. Validation can be split into five formal categories: face validity, internal validity, cross validity, external validity, and predictive validity (Eddy et al, 2012).

Face validity is a subjective assessment conducted by impartial experts, who may evaluate parts of the model (e.g. data sources, assumptions) or the model as a whole, including results. Internal validity, or verification, on the other hand, may be conducted by the original author and includes assessing the mathematical accuracy of the model (Eddy et al, 2012). Checklists such as the TECH-VER (Buyukkaramikli et al, 2018; See Appendix 5) may be used to guide the verification process. Cross, external, and predictive validities are all assessed by comparison. Evaluating cross validity includes comparison with predictions from different models, whereas external validity compares model estimates with actual event data. Finally,

predictive validity evaluates the model's ability to predict study outcomes before they are observed (Eddy et al, 2012).

3.6 Economic Evaluation in Finnish Health Care

In Finland, marketing authorization for new treatments can be sought from a centralized procedure through the European Medicines Agency (EMA), or by a national procedure through the Finnish Medicines Agency (Fimea). In order to receive market authorization, the drug must fulfill three criteria: 1) clinical benefits should outweigh potential risks (adverse events); 2) the product needs to fulfill quality requirements set in the pharmacopoeia; and 3) the formulation and other information should be appropriately disclosed (Fimea, 2020). Market authorization is a prerequisite to selling pharmaceutics and does not require an economic evaluation to be submitted. However, in order to achieve real penetration in the market and clinical use, positive reimbursement status may be critical.

The Pharmaceuticals Pricing Board (lääkkeiden hintalautakunta, HILA) is responsible for granting reimbursement status to medicine used in Finnish health care. The applicant (i.e. manufacturer) is required to send a health economic assessment along with multiple attachments, including a report of clinical benefits, costs, and estimated patient population. Special reimbursement status can be applied for after basic reimbursement has been granted (HILA, 2020). HILA is an agency operating under the Ministry of Health Affairs and Health. The guidelines for economic evaluations submitted alongside relevant applications are set and coordinated by Fimea and are based on the Health Technology Assessment Core Model created by the European Network for Health Technology Assessment (EUnetHTA) (Fimea, 2012).

4. Research Methods

4.1 Economic Evaluation

4.1.1 Type of Analysis

This thesis conducts a cost-utility analysis to determine the incremental cost of providing an additional quality-adjusted life-year using vortioxetine as opposed to escitalopram, venlafaxine, or placebo in first-line treatment of moderate to severe major depressive disorder in Finland. The method is consistent with Finnish guidelines on conducting economic analysis on pharmaceutical products (Fimea, 2012).

4.1.2 Population

The population considered is a cohort representing patients with moderate to severe major depressive disorder in Finland between the ages of 18 and 64. The age restriction is due to data availability, as well as clinical practice, as both the young and the elderly have their own specific treatment considerations (Duodecim, 2020). In addition, restricting the population to adults of working age allows for simplicity in estimating productivity costs from a societal perspective. Severity is assessed as a MADRS score of 30 or above and it is assumed the patients have not received previous pharmacological treatment for their condition. Finally, it is assumed that the patients do not suffer from other psychiatric disorders that may affect treatment effectiveness or overall quality of life.

4.1.3 Intervention

The intervention treatment vortioxetine has been identified due to its relative novelty on the market (Ikäheimo, 2014), as well as its possible positive effects on relapse prevention (Ikäheimo, 2014, McIntyre, 2017). In addition, Finnish guidelines focus on the use of vortioxetine in cases of treatment-resistance, but studies elsewhere have demonstrated positive first-line effects (McIntyre, 2017). Treatment with vortioxetine is started with a 10mg daily dose and can vary between 5-20mg per day during the treatment process (Duodecim, 2020). Due to data availability and to restrict the number of assumptions needed regarding the patient population, this study does not make a distinction between patients on different dose regimens. Treatment response in clinical practice is critically assessed since week 6, and treatment is continued for as long as it is considered effective, until the patient is completely cured. This entails the follow-up and maintenance periods, as described in Section 2.5.

4.1.4 Comparators

The comparators escitalopram and venlafaxine are representative of the current standards of care, as they are the most prescribed antidepressants on the Finnish market (Fimea & KELA, 2018). The pharmacological mechanisms of the individual treatments have been described in Section 2.5. Escitalopram treatment in Finland is initiated with a daily dose of 10mg, which can, with the clinician's discretion, be raised to 20mg per day. Venlafaxine, on the other hand, is started with 75mg per day, and dose regimens can be up to 375mg per day.

As with the intervention treatment, no assumptions about dosage are made in this study regarding the comparators and the same treatment algorithm applies, including follow-up and maintenance phases. In addition to the active treatment comparators, a placebo comparison is included for improved validity, as most clinical data is presented with respect to placebo.

4.1.4.1 Subsequent Treatment Arms

Pharmaceutical treatment of depression is associated with a high probability of treatment switch due to lack of response and/or adverse events (Heiskanen et al, 2011). This study assumes that subsequent treatment lines consist of the same three treatment options presented in first-line treatment. This assumption is due to clinical relevance, as the use of SSRIs (escitalopram) and SNRIs (venlafaxine) is equivalently common in second-line treatment, as they are as first-line treatment options (Duodecim, 2020). In clinical practice today, most patients would receive first-line SSRI or SNRI, and switch to the other in the case of a treatment switch. In case of treatment-resistant depression (3rd line), the same treatment alternatives are included as options. Other antidepressants are not considered due to increasing complexity of the model and presumed lack of significance for end results.

4.1.5 Outcome Measures

The primary outcome of the analysis is the ICER, representing the incremental cost per QALY gained. As the ICER is an incremental measure, it allows for comparison between treatment alternatives. Secondary outcomes include QALYs and life years (lost to suicide) to determine the absolute clinical benefit of each treatment alternative. These are established by calculating the amount of time spent in each respective health state, including death. In addition, the effect of uncertainty is measured with scenario analyses and a probabilistic sensitivity analysis and examined with cost-effectiveness acceptability curves (CEAC) and value of information techniques, including the expected value of perfect information (EVPI).

4.1.6 Time Horizon

The base analysis includes a time horizon of 20 years. Although requiring prolonged extrapolation of data and assumptions regarding the patients' futures, the recurring nature of the condition requires a long-term horizon.

Given that the average age of the patients enrolled in many clinical trials is just over 40 (see Appendix 2), 20 years is considered the longest possible time horizon while still assuming most of the patients remain 64 years old or younger. Sensitivity analyses provide results from 1-, 5-, 10, and 25-year horizons to see the effect of extrapolation. Costs and effects occurring in the future are discounted at an even annual rate of 3%, as advised by the Finnish Medicines Agency (2012).

4.1.7 Perspective

The analysis is conducted from a healthcare perspective, accounting for all depression-related costs falling directly on the budget of health care and social services. This is in line with Finnish guidelines (2012). A scenario analysis also considers a societal perspective, which includes productivity losses arising from absenteeism (sick days) and early retirement (disability pensions). Considering the effect of productivity costs is especially important in the context of depression, given its role as a major cause of early retirement (Finnish Centre for Pensions, 2020). Studies from, for example, Sweden (Ekman et al, 2015) have demonstrated that indirect costs due to productivity losses may account for the vast majority (88% in their study) of costs arising from depression, when assessed from a societal perspective.

4.1.8 Model Structure

The model built for this thesis is a cohort Markov model built on Microsoft Excel 2019. Patients enter first-line acute treatment with MDD and receive one of the following treatments: 5-20mg of vortioxetine; 10-20mg escitalopram; 75-375mg venlafaxine; or placebo. Patients continue with their first-line treatment if they show a response, defined as 50% decrease in the score of a relevant rating scale (studies used MADRS and HAM-D). If patients achieve remission, defined as a score below a certain threshold (12 for MADRS, 7 for HAM-D), they will move to a remissive health state. The efficacy thresholds are in line with the primary data source of Cipriani et al (2018). In accordance with Finnish treatment guidelines, patients are considered recovered after they have stayed in the remissive health state for six months. After recovery, no risk of relapse is assumed.

There are three pathways for patients to enter subsequent treatment lines. During the remissive state, patients are at a risk of relapsing. If they experience a relapse, they will move to second-line treatment. Alternatively, if patients do not respond to first-line treatment within the first 8 weeks, they will switch directly

to second line treatment. This will also occur if patients respond to treatment but fail to achieve remission within 6 months. Finally, each treatment is associated with a risk of treatment switch-inducing adverse events. If patients experience these events, they will switch to a subsequent line of treatment.

Treatment-resistant depression is defined as a lack of response to at least two previous lines of treatment. This model includes a state encompassing all lines of treatment including and after 3rd line. These treatment lines have been consolidated due to the lack of data and consequent uncertainty associated with later treatment lines. In addition, decisions made with recurrent depression are expected to have little impact on first-line decisions and are discounted, given that they occur in the future. For patients experiencing treatment-resistant depression, recovery is achieved after one year in a remissive state. This reflects clinical practice, where a longer maintenance period is applied for difficult-to-treat cases.

A cohort of 1000 hypothetical patients started at cycle 0 in each of the treatment alternatives under investigation. The model utilizes a cycle time of eight weeks to reflect the assessment period used in the data (Cipriani et al, 2018), as well as clinical practice (Duodecim, 2020). Seven health states are included: Depressed (i.e. first-line treatment), remission, 2nd line treatment, 2nd remission, treatment-resistant depression, 3rd remission, and recovered. These health states are considered relevant to the research question and representative of a patient's clinical pathway through the pharmacological treatment of depression. Each health state is associated with a probability of adverse event incidence, which is assumed to occur during the first cycle of initiating a new treatment. In addition, a risk of suicide mortality is associated with both the remissive health state (low risk) and the depressed, active treatment health states (high risk). All patients start from a depressed state. The model structure is illustrated in Figure 3 below.

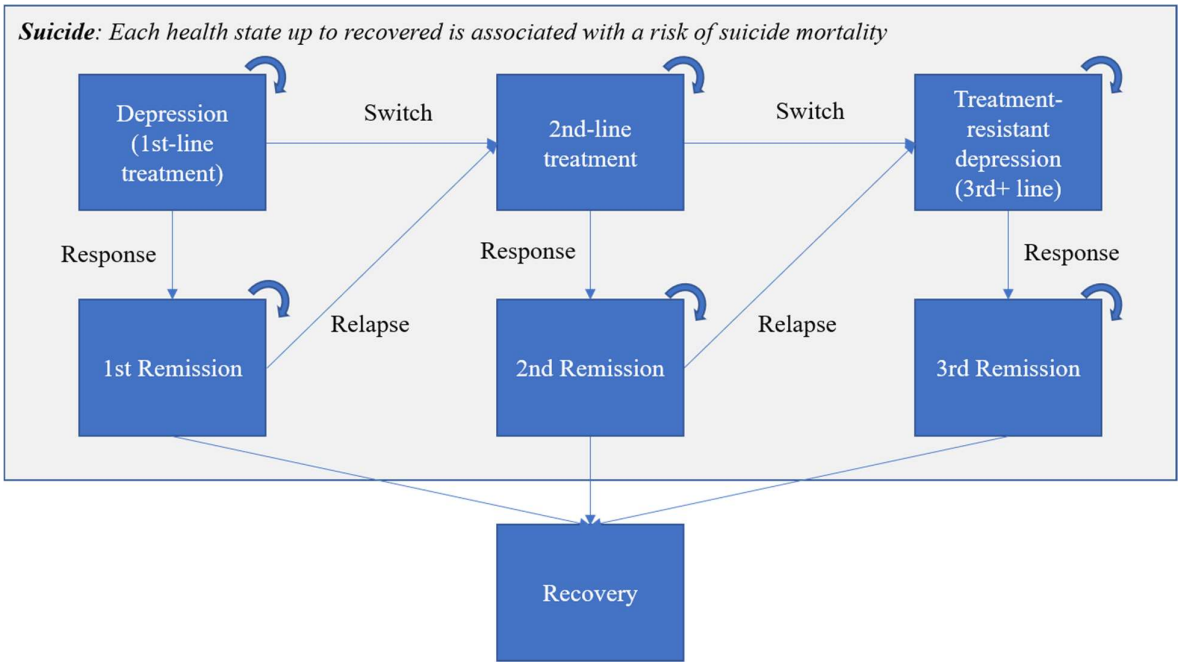


Figure 3: Graphical representation of the Markov chain structure. Each blue rectangle represents a health-state, and arrows between them portray pathways in which patients can move from one cycle to the next in the model. All patients start from the top-left health state (depression). A "switch" includes treatment switches for all reasons, including lack of response and severe adverse events. All health states within the grey rectangle are subject to a risk of suicide mortality, and recovery is an absorbing health state (i.e. once entered, patients remain in that state). Each health state is associated with the possibility of remaining in that respective health state from one cycle to another, as indicated by the arrows in the top right corner of each rectangle.

4.2 Model Parameters

4.2.1 Systematic Literature Review

4.2.1.1 Eligibility Criteria

Studies included in the systematic literature review are RCTs assessing the efficacy and safety of vortioxetine, escitalopram, and venlafaxine against each other and/or placebo in the acute treatment of moderate to severe MDD. Eligible trials should enroll adult patients with a primary diagnosis of MDD and should not explicitly search for patients with (mental) comorbidities, such as anxiety of cognitive dysfunction. Trials are expected to include a population with moderate to severe depression as indicated by a baseline MADRS or HAM-D score of 30 or 18, respectively. Eligible studies should include an outcome measure of remission as defined by a MADRS or HAM-D score under the thresholds of 12 or 7, respectively. No restrictions are imposed based on sex, study location, or inpatient versus outpatient treatment. Studies (or treatment arms of studies) are

excluded if the dosage is not in accordance with Finnish guidelines for vortioxetine (5-20mg/day), escitalopram (10-20mg/day), or venlafaxine (75-375mg/day), respectively. The literature review includes studies published in the English or Finnish languages.

4.2.1.2 Literature Search

The literature review includes studies published in the following databases up to and including March 1st, 2020: PubMed, EMBASE, Clinicaltrials.gov, and the Wiley Online Library. Search terms and algorithms can be found in Appendix 1 of this thesis.

4.2.1.3 Results

The literature review resulted in 796 identified records, out of which 732 were excluded based on title. 64 records were assessed for eligibility based on the full-text or abstract. A further 49 were excluded for reasons listed in Figure 4 below. 15 RCTs were identified that met the study criteria:

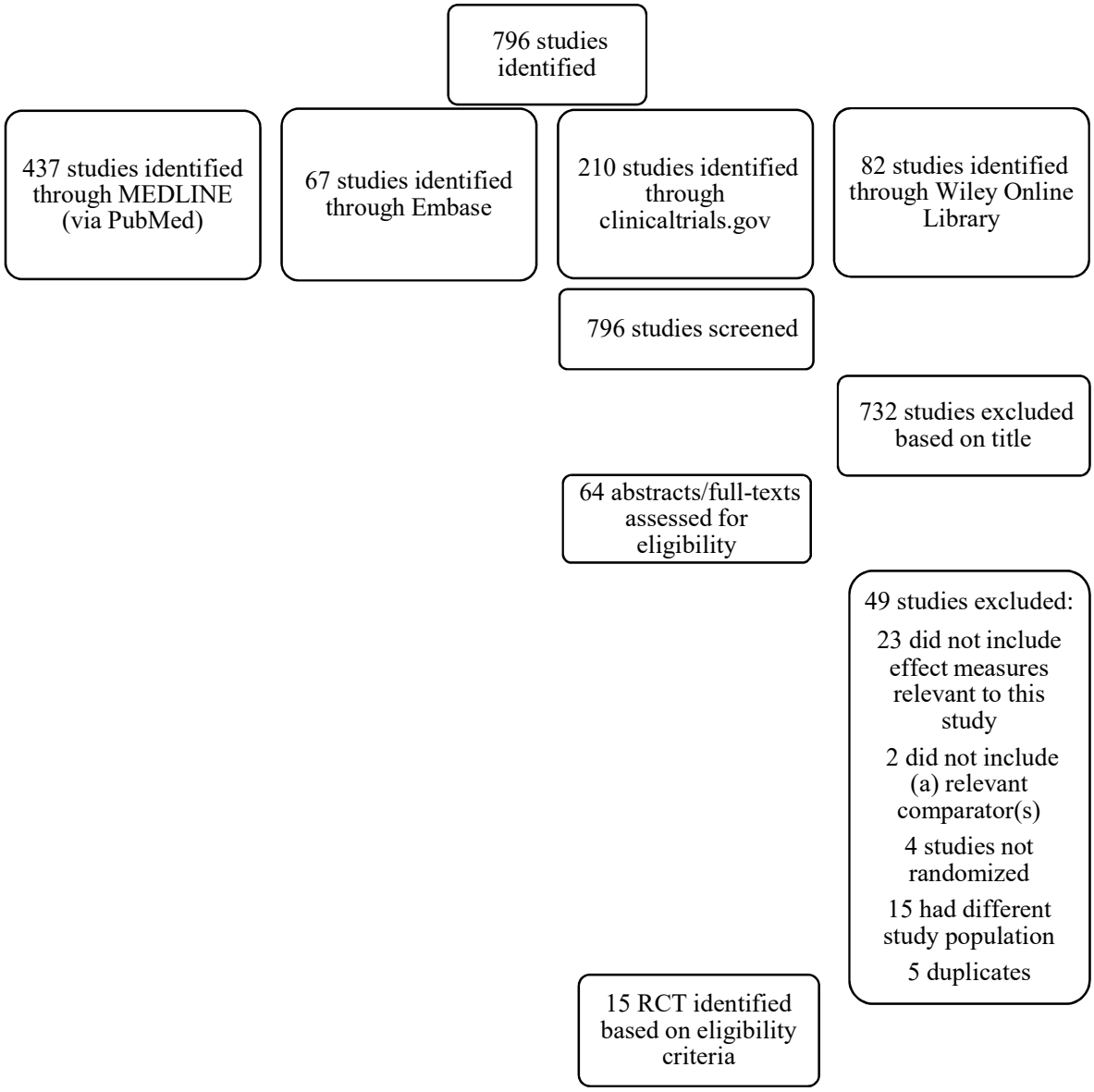


Figure 4: Study Flow Diagram

Key characteristics of the studies identified by the systematic literature are presented in Appendix 2. The literature review also identified a comprehensive systematic review and meta-analysis by Cipriani et al (2018), which was identified as the primary data source for this analysis. The study includes a random-effects network meta-analysis within a Bayesian framework. It was published in 2018 and includes all the treatment arms included in this thesis, as well as nearly all the studies identified by the systematic literature review, and unpublished data retrieved from companies. The literature review presented in the paper was conducted independently by six pairs of investigators. Due to the comprehensiveness, relevance, and applicability of the

analysis by Cipriani et al, a separate (network) meta-analysis was not conducted for the purposes of this thesis. Rather, the results by Cipriani et al (2018) were supplemented by studies identified in the literature review as needed for the purposes of this thesis.

4.2.1.4 Structured Search

The literature review allowed for the identification of additional systematic reviews by Baldwin et al (2007, 2016) and Jakobsen et al (2017), which were used in favor of single RCTs in order to avoid biases arising from using a single source. In order to complement the RCT data, a structured search was conducted for existing cost-effectiveness studies, as well as for relevant cost data and health-related utility values. The structured search used the Tufts Medical Centre CEA registry for existing cost-effectiveness analyses, and Google Scholar for cost-effectiveness, cost data, and utility values. Due to the relative scarcity of existing studies, an extensive literature review was not considered necessary. In addition, cost data is made available through the Finnish Institute of Health and Welfare (THL) and the Social Insurance Institute (KELA). Utility values in the most relevant previous study corresponded with the health states in the current model and were adjusted for Finnish values. Therefore, additional research into utility values was not considered necessary.

The parameters of the model identified through the structure search and the literature review are presented in Table 2 below, along with the respective sources. Where possible, the same data source was used for multiple parameters in order to achieve internal validity and consistency across the Markov model. In addition, to increase transparency, an extensive list of parameters and their values has been presented in Appendix 3.

Table 2: Data sources identified in the systematic review and used in estimating the key model parameters. The table presents the respective study used as a source for each parameter, as well as how that parameter is presented in the source material, and the type of the original study.
HRQoL: Health-Related Quality of Life; RCT: Randomized Controlled Trial; SR: Systematic Review

Parameter	**Study**	**Presentation in original study**	**Type of study**
Efficacy**: Response and remission rates, vortioxetine, escitalopram, venlafaxine** Safety: **AE-related dropout rates, vortioxetine, escitalopram, venlafaxine**	Cipriani et al (2018)	Odds ratios with respect to placebo	SR & Network meta-analysis
Efficacy: **Placebo response rate**	Furukawa & Cipriani et al (2016)	Percentage of patients responding to placebo	SR & Meta-analysis
Efficacy: **Placebo remission rate**	Jakobsen et al (2017)	Number remitting in placebo treatment	
Efficacy: **Recurrent depression, remission rate**	Rush et al (2006)	Percentage of patients responding to 3rd line treatment	Open label interventional trial (STAR*D)
Relapse rate**: Vortioxetine**	Boulenger et al (2012)	Risk ratio relative to placebo	Double-blind RCT
Relapse rate**: Placebo**	Boulenger et al (2012)	Percentage of patient relapsing during placebo treatment	
Relapse rate**: Escitalopram**	Rapaport et al (2004)	Risk ratio relative to placebo	
Relapse rate**: Venlafaxine**	Simon et al (2004)	Odds ratio relative to placebo	
Relapse rate**: 2nd line, 3rd line**	Soini et al (2017), NICE TA 367	Percentage relapsing during remissive period	Cost-utility analysis
Recovery rate: **Recurrent depression**	Markkula et al (2016)	Percentage of depressed still carrying diagnosis after 11 years in Finland	Longitudinal survey-based study
Effects**: HRQoL utilities**	Soini et al (2017)	HRQoL utility score	Cost-utility analysis
Effects**: Disutility associated with AEs**	Soini et al (2017)	HRQoL utility score	
Costs: **Drug acquisition costs**	KELA (2020)	Cost per package	Official Finnish database
Costs: **Resource use costs**	Kapiainen et al (2014)	Cost per unit in 2011	Report for THL
Safety**: Dropouts due to AEs, placebo**	Baldwin et al (2016)	Percentage dropping out due to AEs	Review of RCTs and open-label extension studies
Safety: **AE incidence, vortioxetine, placebo**	Baldwin et al (2016)	Percentage experiencing AEs	
Safety: **AE incidence, escitalopram**	Baldwin et al (2007)	Percentage experiencing AEs	
Safety: **AE incidence, venlafaxine**	Baldwin et al (2007), Baldwin et al (2016)	Percentage experiencing AEs	

4.2.2 Treatment Efficacy

As mentioned in the introduction, the systematic review and network meta-analysis conducted by Cipriani et al (2018) is referenced in the Finnish treatment guidelines for depression (2020) as the most comprehensive meta-analysis of antidepressants to date. The analysis includes 522 double-blind randomized controlled trials including 116 477 patients and comparing 21 different antidepressants to either placebo (304 studies) or each other. 86 of the studies were unpublished at the time. The results of the analysis indicate that all antidepressants are more effective than placebo during an 8-week treatment period. According to the Finnish Physicians' Association, these results are transferrable to the Finnish context.

The study by Cipriani et al (2018) focused exclusively on acute care on an adult population with a primary diagnosis of major depressive disorder. It excluded trials including 20% or more of participants with significant comorbidities, such as bipolar disorder, psychotic depression, or treatment-resistant depression. Cipriani et al (2018) assessed the individual studies for bias according to the Cochrane Handbook for Systematic Reviews of Interventions. They conducted a random-effects meta-analysis to estimate summary odds ratios for the primary outcomes: response rate and treatment discontinuation. Remission rates and discontinuation due to adverse events were included in the supplementary material. For the purposes of this thesis, the odds ratios were translated to risk ratios using the following formula (Zhang & Yu, 1998):

$$RR = \frac{OR}{(1 - P_0) + (P_0 \times OR)} \tag{7}$$

From risk ratios, the transition probabilities describing the efficacy of each treatment line were estimated with regards to placebo by a simple multiplication of the respective risk ratio and the placebo transition probability. Cipriani et al (2018) do not present transition probabilities in their network meta-analysis. Therefore, placebo effectiveness has been retrieved from another study co-authored by Cipriani (Furukawa et al, 2016) (response) and from a meta-analysis comparing SSRI remission with placebo (Jakobsen et al, 2017) (remission). Both studies are methodologically in line with this thesis. A summary of the results is presented in Table 3:

Table 3: Summary of effectiveness parameters applied in each cycle throughout the model.

Treatment	Odds Ratio		Risk ratio		Probability		Standard error	
	Response	Remission	Response	Remission	Response	Remission	Response	Remission
Vortioxetine	1.66	1.49	1.34	1.32	0.48	0.34	0.07	0.07
Escitalopram	1.68	1.64	1.35	1.41	0.49	0.36	0.06	0.06
Venlafaxine	1.78	1.70	1.39	1.44	0.50	0.37	0.05	0.05
Placebo					0.36	0.25	0.02	0.01

4.2.2.1 Subsequent Treatment Lines

Second-line treatment efficacy has been estimated with the method recommended by the National Institute of Care Excellence (NICE) in the technology assessment (TA367) attached to the manufacturer's submission of cost-effectiveness to the Institute. In TA367, the evidence review group (ERG) recommended a proportional reduction in effectiveness according to the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial. The STAR*D trial provides effectiveness measures through four lines of treatment and, regarding the treatment arms currently under study, included venlafaxine. Therefore, the proportional reduction from one treatment line to the next has been taken as a guideline for second- and third-line effectiveness.

Table 4: Subsequent treatment line efficacy applied in each cycle throughout the model

Treatment	Response		Remission		
	1L	2L	1L	2L	3L
STAR*D	0.49	0.28	0.37	0.25	0.11
Vortioxetine	0.48	0.28	0.34	0.23	0.10
Escitalopram	0.49	0.28	0.36	0.24	
Venlafaxine	0.50	0.29	0.37	0.25	

As presented in Table 4, differences in efficacy are mitigated in second-line treatment. Therefore, given the small differences and discounting occurring in the future, third-line efficacy is considered uniform across treatment alternatives. A separate response parameter is not necessary for the third-line context, as the model does not explicitly include switches to further lines after lack of response in third-line treatment.

4.2.3 Recovery

Recovery following remission during first- or second-line treatment was assumed to occur after remaining in a remissive state for 6 months. This was estimated directly from the Markov chain in the model and thus no separate parameter was required. For treatment-resistant depression, the longer maintenance period of one-year did not allow for accurate estimation of how many patients have remained in a state of remission for the required period. In addition, the treatment resistant state essentially contains within it all further lines of treatment. Therefore, a longitudinal Finnish study, representative of the whole population, has been used to estimate recovery rate. The study by Markkula et al (2016) found that 16% of those diagnosed with MDD in 2000 still had an MDD diagnosis 11 years later in 2011. Therefore, a recovery parameter for patients in the treatment-resistant stage has been calibrated so that approximately 84% of the initial patient cohort has recovered after 11-years. The value of the recovery parameter has been assumed as 0.3, implying that 30% of patients achieving remission from a treatment-resistant stage stay in remission for a full one-year period.

4.2.4 Treatment Safety (Adverse Events)

Treatment safety was assessed in the study by Cipriani et al (2018) as the proportion of patients dropping out of each respective treatment line. Individual events and events not leading to treatment switch were not included. The data from Cipriani et al (2018) has been used to determine the proportion of patients who cannot tolerate respective treatment lines and thus move on to the next treatment line after the first treatment cycle regardless of response status. These probabilities have been determined from ORs similarly to the efficacy parameters described in the section above.

For specific AE incidence, in order to assess associated disutility and costs, data from two systematic review studies by Baldwin et al (2007, 2016) have been used. These studies have been selected due to their comprehensive scope encompassing relevant RCTs, their consistency in methods, as well as their applicability to this thesis. Both studies consider short-term effects (6/8-weeks) during the acute treatment of MDD. The most common AEs were nausea (21-29%) and headache (13%-19%).

Adverse events are assumed to occur additively at treatment initiation for each first-line treatment. This is considered appropriate due to the possibility of patients switching treatment in case AEs are severe. Nevertheless, treatment switch due to AEs in subsequent treatment lines is not assumed, in order to simplify

the model and remove the necessity to make assumptions of the influence of the physician in clinical practice. A physician may be able to use first-line results to select a more suitable second-line treatment, making the relevance of double-blind clinical data questionable. Costs and disutilities associated with AEs from different treatments are, however, considered in subsequent treatment lines as well.

4.2.5 Treatment Switch and Relapse

Patients are subject to a risk of relapse during the maintenance period in remission. This risk is dependent on the initial treatment alternative. Each treatment is associated with their respective risks, which have been retrieved from RCTs considering maintenance treatment specifically. The data sources present the risk as a ratio (risk or odds) with respect to placebo. For second and third-line treatments, the risk of relapse has been retrieved from the study of Soini et al (2017), as well as NICE TA367 (2018) and is assumed uniform across the treatment lines. Previous studies have concluded a lack of evidence of differences between treatments in relapse prevention in treatment-resistant cases, which is consistent with the evidence collected in this study for first line relapse prevention (risk). All first-line treatments are associated with an 8-10% risk of relapse during the maintenance period. The potential of vortioxetine in relapse prevention alluded to in earlier chapters in this thesis was not apparent in the data gathered during the literature review.

4.2.6 Subsequent Treatment Line Composition

As mentioned in sections above, subsequent treatment lines are assumed to consist of a mixture of the first-line treatment options. However, an uneven mix is assumed due to differences in clinical practice and unequal market share of the products. SSRIs (escitalopram) and SNRIs (venlafaxine) are assumed as preferred second-line treatment options due to their established positions in the market. It then follows that their use in later treatment might be relatively less, due to the high probability that they have already been tested in an earlier treatment phase. This is especially true for escitalopram, given that it is considered a relatively safe option (Duodecim, 2020) and is the most prescribed antidepressant in Finland (KELA, 2019).

Table 5: Subsequent treatment line composition after respective first-line treatment. Columns represent the treatment with which patients start, whereas the rows describe subsequent treatments in possible second (2L) and third (3L) lines. Composition expressed as a percentage of patients initiating each treatment from all patients entering the respective treatment line (i.e. 20% of patients who do not achieve remission from initial escitalopram treatment enter second line vortioxetine treatment).

		1L Treatment		
Treatment Line	**Treatment Alternative**	Vortioxetine	Escitalopram	Venlafaxine
2L Treatment	Vortioxetine		20%	20%
	Escitalopram	50%		80%
	Venlafaxine	50%	80%	
3L Treatment	Vortioxetine		50%	95%
	Escitalopram	33%		5%
	Venlafaxine	67%	50%	

Given that these assumptions are not directly informed by literature, each of these percentages is associated with a 20% (of the initial value) uncertainty in the probabilistic sensitivity analysis. It is also important to recognize that, since the model assumes uniform effectiveness in third-line treatment, third-line treatment composition is only used for the purposes of estimating costs and AE disutilities.

4.2.7 Suicide Mortality

The model captures suicide mortality by assuming a risk of death throughout the model. An average of 18.98 per 100,000 people aged 20-64 suffered death by suicide in the years between 2013 and 2017 in Finland (THL, 2020). This risk is adjusted in the model so that the patients in a depressed state have a 20-fold risk (Osby et al, 2001) of suicide compared to those in remission. Suicide incidence figures are assumed a 20% uncertainty of the mean value. The costs associated with suicides are omitted, due to their uncertainty and the presumed insignificance to the results of the model, given the marginal differences in suicide mortality between the treatments. Patients in the recovered state are assumed no risk of suicide mortality.

The model does not include non-fatal suicide attempts due to lack of data and the changed care requirements of those with suicide attempts (i.e. they would require a different model after the event). This omission could influence the end results, although it is difficult to state the significance of it, due to the uncertainty associated with measuring both the incidence and costs of non-fatal suicide attempts. Considering the small differences between the treatment alternatives in suicide mortality, it is, however, arguably likely

that the inclusion of non-fatal suicides would not change the relative cost-effectiveness of the assessed treatment alternatives. In addition to assumptions related to suicides, it is assumed there is no other source of mortality (i.e. all-cause mortality is omitted) due to the relatively young age of the cohort (below 65) and consequent low risk of death, which is considered insignificant for the results of the model.

4.2.8 Costs

The model includes four types of costs, namely: drug acquisition costs, adverse events management costs, resource use costs, and productivity costs. Drug costs are straightforward to retrieve, as dosage is not dependent on the individual patient (e.g. body surface) and prices are published by the Social Insurance Institution (KELA). In addition, companies are required by law to provide the same price to all pharmacies in Finland. Therefore, no uncertainty has been assumed for drug acquisition costs, although the effect of price fluctuations is evaluated through a sensitivity analysis.

Price has been taken according to the daily defined dose for each drug (10mg vortioxetine, 10mg escitalopram, 100mg venlafaxine) (KELA, 2019). In addition, the study considers the least costly generic options of escitalopram (Escitalopram Actavis) and venlafaxine (Venlafaxin Orion). Vortioxetine is patent protected, so the price of Brintellix® has been taken as the market price in Finland. Drug acquisition costs are presented in Table 6 below.

Table 6: Drug acquisition information, including specific product and prices per package, dose, and model cycle

Treatment	**Product Name**	**Price per package (€)**	**Price per dose (€)**	**Cost per cycle (€)**
Vortioxetine	Brintellix	111.74 (98 pills, 10mg)	1.14	63.85
Escitalopram	Escitalopram Actavis	4.39 (98 pills, 10mg)	0.04	2.51
Venlafaxine	Venlafaxin Orion	46.29 (98 pills, 150mg)	0.31	17.63

Resource use costs have been estimated according to the figures presented by Soini et al (2017). To the author's knowledge, this is the most recent cost-utility analysis covering MDD in the Finnish context and is representative of the current care practices. The analysis by Soini et al (2017) also includes depressed, remissive, and relapsed health states, although in a slightly different context, as it considers second-line care initiation as the starting point of the analysis.

Table 7: Resource utilization as presented in Soini et al (2017)

Resource Category	**Utilization Per Cycle**			**% Utilization**		
	Depressed	Remission	Relapse	Depressed	Remission	Relapse
GP visits	4.88	0.35	0.31	95%	100%	100%
Psychiatrist visit (60min)	5.88	-0.15	0.39	10%	100%	100%
Psychotherapy or counseling (60min)	3	0.06	0.45	25%	100%	100%
Psychiatric ward (per day)	8	0.9	0.51	1%	100%	100%
Outpatient hospitalization (per day)	0	0.12	0.07	0%	100%	100%
Absenteeism (days)	43	0.94	13.49	27%	100%	100%

Soini et al (2017) take a one-year perspective, which may explain the low resource use of the relapsed category. In addition, the negative resource use of psychiatric visits (-0.15) for patients in remission is explained by the authors as negative cross-elasticity: those using primary care resources cannot use specialized care resources simultaneously. These figures have been adjusted for the present study by assuming that patients in a relapsed state (i.e. second- or third-line active treatment) use the resources associated with both the depressed and the relapse categories presented in Table 7 above, the only exception being the productivity costs associated with absenteeism.

Adverse events are each associated with a respective resource use cost, which is independent from the treatment line the person is in. The management strategies of different AEs have been assumed by severity of the event. For example, symptoms of constipation are assumed to be treated by a simple call to the nurse, whereas sexual dysfunction is managed by a visit to the doctor. Any consequent GP visits resulting from a call to the nurse are assumed to be captured in the resource use variables presented in Table 7 above. Possible pharmaceutical prescriptions for AE management are not expected to have a significant effect on the present model and are therefore not included.

Societal costs are captured by productivity loss due to absenteeism from work, whether due to sick days or early retirement due to depression. The percentage of patient moving to disability pension is guided by Finnish statistics on number of people receiving antidepressants (281 174; THL, 2020) and people moving to disability pension due to depression (3800; Finnish Centre for Pensions, 2020). Adjusted per 8-week cycle, 0.21% of depressed patients move to disability pension each cycle. It is assumed that early retirement happens

only from a treatment-resistant stage. All costs arising from resource use, adverse events, or productivity losses are assumed a 20% uncertainty of the respective mean value.

4.2.9 Utility Values

Utility values required to compute QALYs have been retrieved by the study from Soini et al (2017) for the depressed and remissive health states. These are based on the REVIVE study and adjusted for Finnish preferences. The REVIVE study consisted of adults in second-line treatment of depression (Soini et al, 2017). However, the utility values are not expected to be significantly different for a first-line population, especially given the severe nature of the condition. More importantly, the values presented by Soini et al (2017) are adjusted for Finnish preferences and accurately reflect the health states included in the current model. Utility in the treatment-resistant condition is assumed same as baseline depression.

Disutilities associated with adverse events are also according to the study by Soini et al (2017). Missing values have been assumed by the author according to condition severity (e.g. constipation is assumed equivalent disutility to diarrhea). The original data sources are the REVIVE study, and a study by Cipriani et al (2010). All utility values are assumed a 20% standard error to account for uncertainty.

Table 8: HRQoL Utility Values for health states and adverse events

Event/State	**Utility value**	**Standard error**
Depression (regardless of treatment line)	0.51	0.102
Remission	0.84	0.168
Adverse Event Disutilities		
Constipation	0.10	0.020
Diarrhea	0.10	0.020
Dizziness	0.10	0.020
Dry Mouth	0.10	0.020
Fatigue	0.08	0.016
Headache	0.08	0.016
Hyperhidrosis	0.10	0.020
Insomnia	0.08	0.016
Nausea	0.10	0.020
Sexual dysfunction	0.13	0.026
Somnolence	0.12	0.024
Nasopharyngitis	0.08	0.016

4.3 Uncertainty and Sensitivity Analyses

Structural uncertainty in the model was assessed through scenario analyses, including different time horizons between 1 and 25 years, switching between a societal and a healthcare perspective, varying discount rates, and varying drug acquisition costs. Parameter uncertainty was addressed by a probabilistic sensitivity analysis (PSA), for which appropriate distributions were attributed to each parameter of uncertainty: lognormal distribution for ratios (i.e. odds and risk ratios), beta distributions for probabilities (e.g. placebo effectiveness) and utility values, and gamma distributions for costs (Briggs et al, 2011). Uncertainty estimates were taken from the literature where available or assumed a standard 20% of the parameter's mean value. The parameters where this assumption is applied are presented in Appendix 3. A Monte Carlo simulation was used to draw 5000 samples according to the distributions and calculating the resulting 5000 ICERs using equation (3). 5000 sample iterations are considered to be sufficient to establish stability in the results and their associated uncertainties. The probabilistic ICERs were presented on a CE-plane. Finally, the simulation values were used to calculate incremental net monetary benefits and construct CEACs for all alternatives.

4.3.1 Value of Information (VOI)

EVPI has been assessed with the method presented by Strong et al (2014) using the Sheffield Accelerated Value of Information tool online. This requires using the PSA techniques for sampling individual parameters and uploading them to the website along with the associated costs and effects. The tool uses R software to calculate VOI-measures and the user can customize the specifications of the model (population, time horizon, threshold). For EVPPI, values regarding single parameters have not been analyzed due to the large number of parameters and their interrelatedness. Instead, parameters have been grouped in relevant categories, including utilities, costs, efficacy, and adverse events. Population EVPI was estimated with the modeled population of 1000 patients annually over the time horizon of the model, although real-life implications of this figure are discussed in the discussion section. Finally, the Monte Carlo method has been used to compare results between the methods and discuss their robustness.

5. Results

5.1 Economic Evaluation

5.1.1. Total Costs and Effects

Over a time horizon of 20 years, venlafaxine is associated with the lowest costs per patient from a health care perspective. The costliest first-line treatment according to the model is vortioxetine, although cost of placebo treatment is significantly higher due to the prolonged time spent in a depressive health state and the consequent cumulation of resource use costs. Venlafaxine is expected to be the most effective treatment, while vortioxetine and escitalopram have almost identical effects on QALYs (0.002 difference in favor of vortioxetine). The undiscounted and discounted cost figures per patient are presented in Table 9 below:

Table 9: Total Costs and QALYs of According to Treatment Alternative

Undiscounted			
Treatment	**Total Costs (€)**	**Total QALYs**	**Average Annual Costs (€)**
Venlafaxine	28 255	18.151	1 417
Escitalopram	28 394	18.143	1 424
Vortioxetine	28 507	18.145	1 429
Placebo	31 332	17.897	1 571
Discounted (3%)			
Treatment	**Total Costs (€)**	**Total QALYs**	**Average Annual Costs (€)**
Venlafaxine	24 328	13.518	1 220
Escitalopram	24 441	13.511	1 225
Vortioxetine	24 553	13.513	1 231
Placebo	26 940	13.303	1 351

5.1.2 Disaggregated Costs

The largest cost category is healthcare resource use costs, which accounts for up to 96% of total costs over a 20-year time horizon from the healthcare perspective. Drug acquisition costs were the second largest category, followed by adverse event management, which is a relatively small cost for all treatment alternatives. Nevertheless, venlafaxine is associated with the highest AE management costs, while vortioxetine is expected to incur the highest drug acquisition and healthcare resource use costs.

Table 10: Disaggregated Costs per Patient According to Treatment Alternative

Disaggregated Costs (per patient, €)			
Treatment	Drug acquisition costs	Healthcare resource use costs	Adverse event management costs
Venlafaxine	1 107	27 551	64
Escitalopram	1 096	27 708	58
Vortioxetine	1 206	27 729	50
Placebo	-	31 795	29

5.1.3 Life Years Lost to Suicide

Over 20 years, patients starting on escitalopram are expected to lose the most QALYs due to depression. However, in terms of life years, vortioxetine is associated with the highest losses, although only by a slight margin. In the base case analysis with a cohort of 1000 patients, those starting on vortioxetine lose a total of 8.55 life years due to suicide, whereas the respective numbers for escitalopram and venlafaxine are 8.53 and 8.48. Nevertheless, all active treatments are associated with the same probability of 0.06% for dying of suicide. For placebo, life years lost to suicide amount to 9.79 with a probability of 0.07% of suicide mortality. The difference between placebo and the active treatments may demonstrate the opportunity for suicide prevention if patients receive the appropriate care.

5.1.4 Cost-Effectiveness

Table 11 below is an adaptation from *Methods for the Economic Evaluation of Health Care Programs* by Drummond et al (2015). It is a way of presenting ICERs and net benefits with multiple alternatives and illustrate the need to consider all available alternatives when making decisions. In Table 11, vortioxetine is associated with an ICER of €54 893 when compared to escitalopram (discounted results), which may seem like a viable option given a sufficiently high threshold. However, this comparison is only relevant when vortioxetine and escitalopram are the only options available. When considering venlafaxine as well, it becomes clear that vortioxetine is not the optimal choice according to the analysis, as venlafaxine has both higher benefits and lower costs, indicated by SD in the columns of vortioxetine: strongly dominated. As stated by Drummond et al (2015), this illustrates why strongly dominated alternatives should not be considered as a basis of ICER comparisons, which is also why the relevant alternative for vortioxetine is venlafaxine rather than escitalopram. Placebo is associated with higher costs and smaller benefits than each of the active treatment

alternatives and is effectively dominated by all of them. Therefore, placebo is never considered a cost-effective alternative. A key feature of Table 11 is the organization of the treatment alternatives in order of ascending costs. This allows for ease of interpretation, as the table can be read from the top down.

Table 11: Deterministic *ICERs over a 20-year time horizon*

Undiscounted							
Treatment alternative	**Cost (€)**	**QALYs**	**ICER compared to (€)**			**Net benefit (€)**	
			Lowest Cost	**Next lowest cost**	**Relevant alternative**	**€30 000 per QALY**	**€60 000 per QALY**
Venlafaxine	28 255	18.152	-	-	-	516 284	1 060 824
Escitalopram	28 394	18.143	SD	SD	SD	515 907	1 060 208
Vortioxetine	28 507	18.145	SD	58 124	SD	515 852	1 060 212
Placebo	31 332	17.897	SD	SD	SD	505 566	1 042 465
Discounted (3%)							
Treatment alternative	**Cost (€)**	**QALYs**	**ICER compared to (€)**			**Net benefit (€)**	
			Lowest Cost	**Next lowest cost**	**Relevant alternative**	**€30 000 per QALY**	**€60 000 per QALY**
Venlafaxine	24 328	13.518	-	-	-	381 202	786 731
Escitalopram	24 441	13.511	SD	SD	SD	380 896	786 233
Vortioxetine	24 553	13.513	SD	54 893	SD	381 845	786 243
Placebo	26 940	13.303	SD	SD	SD	372 147	771 234

*Strongly dominated (SD) means that there is another treatment with higher expected benefits and lower costs than the respective alternative; (-) indicates no comparison: in this case, venlafaxine would be compared with itself. Example of net benefit estimation: Discounted net benefit of venlafaxine for €30 000 threshold = €30 000 * 13.5176 - €24 328 = €381 202. Adapted from Drummond et al (2015)*

The net benefit figures illustrate the effect of a threshold value when considering cost-effectiveness. As there is no definitive threshold value employed in Finland, Table 11 presents two values for illustrative purposes. €30 000 or equivalent is a common threshold in other countries, such as the UK, and €60 000 is taken as an illustration of a relatively high threshold value. From the comparative figures, it is apparent that the relationship

between the net benefits of each respective treatment option vary as the threshold increases to values above the ICER.

5.2 Sensitivity Analysis

5.2.1 Scenario Analysis

The following sections present the results of the scenario analyses with regards to perspective, time horizon, discount rates, and drug acquisition costs. For tables including full cost-effectiveness results of the scenario analysis, please refer to Appendix 3.

5.2.1.1. Societal Perspective

The analysis from a societal perspective includes productivity costs from both absenteeism and early retirement. Venlafaxine remained the dominating alternative, and vortioxetine the most expensive (excluding placebo). In a societal perspective, productivity costs accounted for approximately 70% of total costs for all treatment alternatives over 20-years. A maximum of approximately 32 patients were estimated to retire from the workforce early due to depression in each treatment alternative. Total discounted costs per patient were between €78 000 and €80 000 for each active treatment.

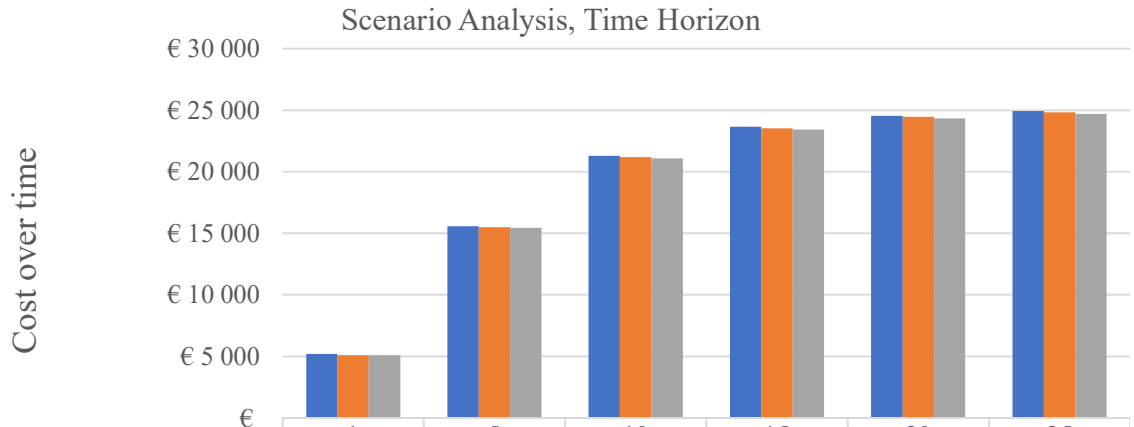
5.2.1.2 Time Horizon

A sensitivity analysis considering a one-year time horizon revealed different results from the base case deterministic analysis. The most affordable (least costly) option was escitalopram, which dominated venlafaxine as a first-line treatment option. Vortioxetine was associated with both higher costs, and greater benefits with an ICER of €44 616 against the relevant alternative, escitalopram. No discounting was applied due to the short time horizon.

Table 12: Scenario analysis considering a one-year time horizon

1-year time perspective					
			ICER (€) compared to		
Treatment	**Cost (€)**	**QALYs**	**Lowest Cost**	**Next lowest cost**	**Relevant alternative**
Escitalopram	5 097	0.701	-	-	-
Venlafaxine	5 105	0.700	SD	SD	SD
Vortioxetine	5 204	0.703	44 616	28 620	44 616
Placebo	5 437	0.676	SD	SD	SD

A 5-year time horizon presents similar conclusions to the base case analysis. Venlafaxine is both the most affordable and most effective treatment option. Total costs are approximately €15 500 for all treatment alternatives, with the difference between the most expensive option (vortioxetine) and venlafaxine being €167. The same relative results prevail over time. Very little difference is observed in costs between the 20-year base case scenario and the 25-year time horizon. This is a reflection of both discounting in the long term, as well as most patients recovering before the 20th year. Figure 5 illustrates the effect of time horizon below:



	1	5	10	15	20	25
Vortioxetine (Cost)	5 204 €	15 592 €	21 289 €	23 643 €	24 553 €	24 929 €
Escitalopram (Cost)	5 097 €	15 483 €	21 179 €	23 531 €	24 441 €	24 817 €
Venlafaxine (Cost)	5 105 €	15 425 €	21 085 €	23 423 €	24 328 €	24 701 €
Vortioxetine (QALYs)	0,703	3,698	7,284	10,640	13,513	16,097
Escitalopram (QALYs)	0,701	3,695	7,281	10,638	13,511	16,095
Venlafaxine (QALYs)	0,700	3,698	7,287	10,644	13,518	16,101

Figure 5: Sensitivity analysis, time horizon

5.2.1.3. Differential Discounting

To assess the effect of discounting and the robustness of the base-case results, a sensitivity analysis with differential discount rates is presented. Costs are discounted at a higher rate of 4%, whereas effects are discounted at a rate of 1.5%. This is standard practice in countries such as the Netherlands and has been argued to better reflect the value of health and capital over time (Attema et al, 2018). However, the adjusted discount rates have minimal effect on cost-effectiveness in the base-case analysis. Venlafaxine remains the dominating alternative, and vortioxetine is associated with an ICER of €54 222 per QALY against escitalopram.

5.2.1.4 Drug Acquisition Cost Analysis

As a test for the effect of drug acquisition costs, a sensitivity analysis is conducted where all treatment alternatives have the same drug acquisition costs as the alternative with the lowest acquisition cost in the base case analysis (escitalopram). This cost is €2.51 per cycle. Venlafaxine remains the optimal treatment with lower costs and greater benefits than the other two active treatment alternatives. Excluding venlafaxine, vortioxetine is associated with an ICER of €4 827 per QALY compared to escitalopram in this scenario.

Finally, as venlafaxine is the least costly treatment overall, the price point at which an increase in the drug acquisition costs of venlafaxine causes a difference in the cost-effectiveness decision is tested. This is done with the built-in goal seek function on Microsoft Excel, by testing at which price points the ICER between vortioxetine and venlafaxine becomes €0 and €30 000, respectively. In order to make venlafaxine treatment equally costly to the most expensive active treatment alternative, vortioxetine, the price per cycle would have to be increased from €17.63 to €182.04. If that were to happen, escitalopram would be the least costly option, but venlafaxine would be considered cost-effective at common thresholds, with an ICER of €10 467 compared to escitalopram. Vortioxetine would remain dominated by venlafaxine, as it would still be associated with lower benefits. The results are presented in Table 13 below:

Table 13: Scenario analysis with an increased price of venlafaxine. Acquisition cost of venlafaxine has been increased to make total costs equivalent to the costliest active treatment alternative, vortioxetine.

Venlafaxine cost increased from €17.63 per cycle to €182.04					
			ICER (€) compared to		
	Cost (€)	**QALYs**	**Lowest Cost**	**Next lowest cost**	**Relevant alternative**
Escitalopram	26 071	13.511	-	-	-
Venlafaxine	26 138	13.518	10 467	10 467	10 467
Vortioxetine	26 138	13.513	32 927	SD	SD
Placebo	26 940	13.303	SD	SD	SD

In order to change the decision at common thresholds, venlafaxine would need to be associated with an ICER of €30 000 or above when compared to the other treatment alternatives. In this case, the acquisition cost per cycle would be raised to €277.51. If the model assumed that venlafaxine would always be used at the maximum approved dose (375mg), acquisition cost per cycle would be €66.13. Given that this value is substantially smaller compared with both €182, as presented in Table 13, and €278 presented in the additional analysis, the

results are arguably robust to price changes occurring from, for example, higher average dosing than assumed by the model.

5.2.2 Probabilistic Sensitivity Analysis

The probabilistic sensitivity analysis (PSA) included 5000 simulations, for which the key summary statistics are shown below in Table 14:

Table 14: PSA summary statistics. 2.5% and 97.5%, respectively, represent the lower and upper boundaries of the credible interval of results.

	Costs (€)				**QALYs**			
	Vortioxetine	Escitalopram	Venlafaxine	Placebo	Vortioxetine	Escitalopram	Venlafaxine	Placebo
Mean	25 409	25 284	25 195	27 116	13.454	13.453	13.458	13.285
Min	8 983	8 845	8 873	9 357	10.814	10.690	10.670	10.326
Max	66 496	56 553	57 025	64 159	14.588	14.576	14.577	14.523
St. dev	5 948	5 923	5 879	6 397	0.468	0.465	0.464	0.516
2.5%	15 449	15 472	15 409	16 491	12.418	12.14	12.425	12.129
97.5%	38 220	38 324	37 954	40 904	14.198	14.193	14.198	14.116

All the pharmacological treatments are associated with a similar level of variance for both costs and QALYs (slight differences in standard deviation). Placebo values for costs display a greater mean value and higher values for both the upper and lower boundaries of the 95% credible interval. In terms of QALYs, all placebo values are lower than the respective values of the antidepressant treatments, except standard deviation. Placebo treatment is associated with a higher variance for both costs and QALYs.

5.2.2.1 Cost-Effectiveness Plane

Results of the PSA may also be presented on a cost-effectiveness plane, where blue diamonds represent the comparison between vortioxetine and escitalopram and orange triangles the comparison between vortioxetine and venlafaxine:

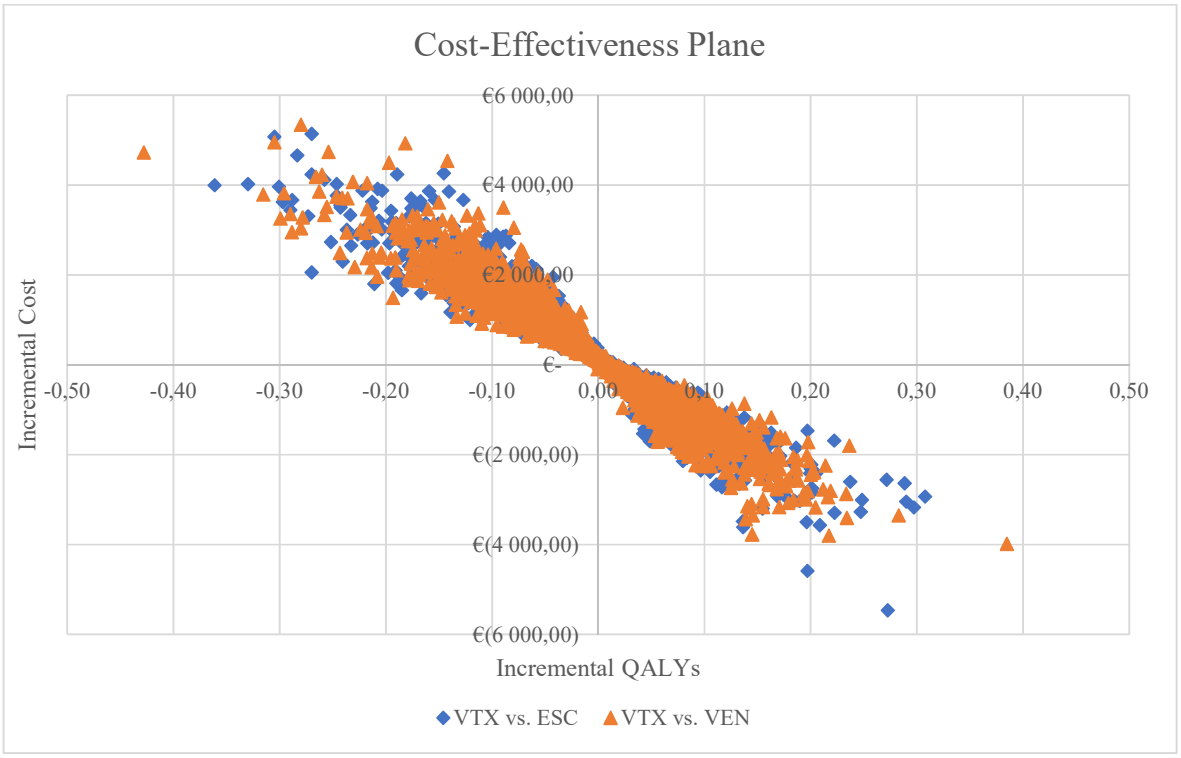


Figure 6: Cost-Effectiveness Plane, Incremental Results. Placebo is not presented, in order to allow easier interpretation of the figure. Most observations against placebo would fall in the southeastern quadrant, given that vortioxetine is dominant over placebo. Two outliers (one from each data series) have been omitted from the figure from the northwest quadrant.

Most of the simulated observations are close to the origin, demonstrating the small differences in costs and efficacy. However, observations are spread in both the northwestern and southeastern quadrants, indicating that the treatment is either dominant or dominated with respect to its comparators. According to the PSA, vortioxetine is never cost-saving at the cost of QALYs (no observations in southwest quadrant), and there is a limited possibility of trade-offs between higher costs and greater QALYs (observations in northeast quadrant very close to origin). The results imply significant uncertainty in making the cost-effective treatment choice, as values with both comparators lie on opposite quadrants.

The relationship between costs and QALYs for each treatment alternative can also be examined by plotting total costs against total QALYs. However, due to small differences between the (active) treatments, it is hard to distinguish any patterns from the scatterplot. Nevertheless, it acts as an illustration of the difference between the active treatments, and placebo. In Figure 7, total costs and QALYs of all treatment alternatives have been plotted on a CE-plane, according to the PSA. As the observations may be hard to distinguish from each other, the horizontal axis has been modified to start from 10.5 QALYs, rather than the origin, and average

observations have been plotted separately with data callouts. From the average observations, it becomes clear that placebo is associated with more costs and less benefits (QALYs) than any of the active treatments. The average points of the active treatments are indistinguishable from each other.

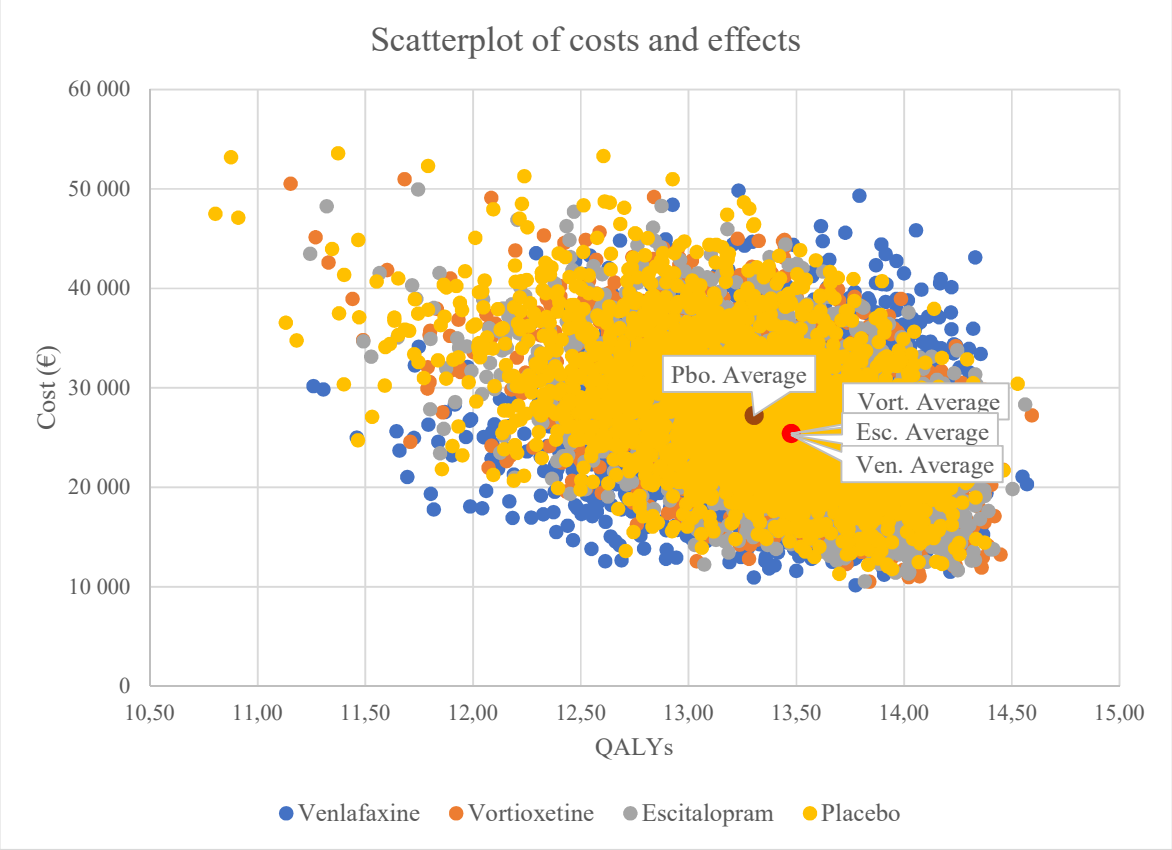


Figure 7: Total costs and QALYs of each simulation of the PSA

5.2.2.2 Cost-Effectiveness Acceptability Curves

The probability of being cost-effective across different thresholds remains relatively constant over threshold values from €0 to €150 000. None of the treatments emerge with a probability above 37% at any threshold. Escitalopram is associated with a probability of cost-effectiveness between 29% and 32%, while vortioxetine varies between 31% and 37%, and venlafaxine is associated with probabilities between 33% and 37% of cost-effectiveness, depending on the threshold value. Venlafaxine is most likely to be cost-effective until a threshold of €25 000, at which it is considered equally likely to be cost-effective as vortioxetine. For thresholds of €30 000 or more, vortioxetine is associated with the highest probability. Placebo is never a cost-effective alternative out of the assessed treatment alternatives, according to the model. Figure 8 illustrates below:

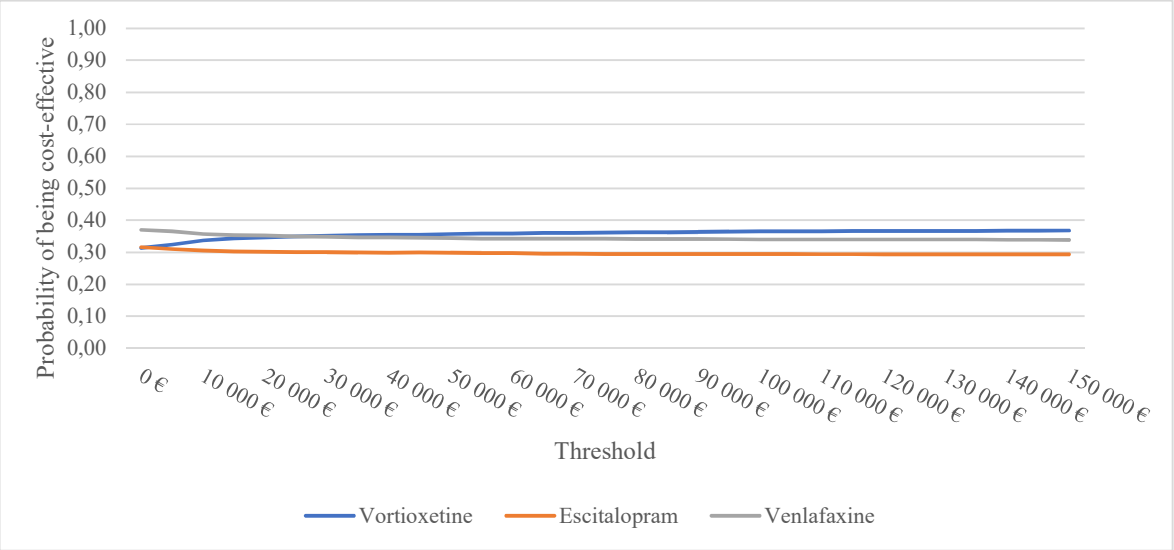


Figure 8: Cost-Effectiveness Acceptability Curves. Placebo is not shown, as the associated probability is 0% across all threshold values.

5.3 Value of Information

5.3.1 Sheffield Accelerated Value of Information

Value of information (VOI) was primarily assessed with the Sheffield Accelerated Value of Information (SAVI) tool online. The CEAC generated by SAVI is very similar to that presented in the previous section, and at a threshold of €20 000, venlafaxine is associated with the highest probability (35.2%) of being cost-effective, although the difference to vortioxetine is marginal. When compared head-to-head with vortioxetine, the SAVI tool estimates that venlafaxine has a slightly higher probability of being cost-effective, 53.9%. Raising the threshold to €30 000 has no significant effect on the results. The results from the SAVI-tool are in accordance with the results presented in sections above, which implies robustness for the EVPI measures.

5.3.2 Expected Value of Perfect Information

Overall EVPI is estimated as €1343 per person affected by the treatment decision at a WTP-threshold of €20 000. For a cohort of 1000 patients over a time horizon of 20 years, this implies that the maximum amount of money that could be spent on additional studies affecting the question would be €26 870 000. Increasing the threshold to €30 000 increases the EVPI to €1733 per person, generating a higher maximum.

5.3.3 Expected Value of Partially Perfect Information

The SAVI tool predicts single parameter EVPPI to be highest for the incidence rate of nasopharyngitis for patients on escitalopram, at €8.98 per patient. However, given the large number of parameters (152), it may be better to group parameters into relevant categories and assess EVPPI accordingly. The assigned categories and the subsequent SAVI-generated analysis is presented in Table 15:

Table 15: EVPPI Results, Threshold €20 000

Category	**Per person EVPPI (€)**	**Annual EVPPI per cohort (1000) (€)**	**EVPPI of cohort per 20 years (€)**
Response rates	3.64	3 637	72 740
Remission rates	14.80	14 802	296 031
Response & Remission rates (efficacy)	6.89	6 886	137 727
Relapse rates	24.06	24 060	481 216
Utilities	17.38	17 384	347 689
2L Treatment sequence	3.26	3 260	65 205
Dropout rates	22.13	22 129	442 574
AE costs	650.57	650 575	13 011 495
AE incidence	939.87	939 869	18 797 383
Resource use costs	65.33	65 328	1 306 558
Suicide rates	0.03	33	660

As seen in Table 15, adverse events incidence stands out as the category with the highest EVPPI: €940 per person. According to these results, more than €18 million could be spent annually to reduce uncertainty in this category. Other parameter categories that stand out include AE costs, and, although with significantly lower EVPPI than AE parameters, resource use costs. Adverse events stand out, as they are a category with multiple parameters (12 for costs, 48 for incidence), each associated with the assumed 20% uncertainty mentioned in sections above. Nevertheless, the results support current clinical practice, where adverse events are a key consideration when choosing an antidepressant (Duodecim, 2020). However, given their relatively small impact on overall costs (see section 5.1), it should be carefully reviewed whether this is the category that is most relevant for further research. Thus, while Table 15 presents some evidence of where further research could be placed, appropriate caution should be practiced when interpreting the results.

5.3.4 Monte Carlo Comparison

The EVPPI results were additionally simulated with the Monte Carlo (MC) method on Microsoft Excel. Authors have shown previously (Tuffaha et al, 2016), that the nonparametric regression approach employed by SAVI can provide equivalent results as the "traditional" Monte Carlo method, with significantly reduced computational requirements. The computational burden imposed by the MC simulation was also apparent in conducting this thesis, allowing for a maximum of 400 outer and 100 inner loops on three different parameter categories. The categories for the MC simulation were chosen based on the results of the SAVI analysis (categories with highest EVPPI). Results with different inner and outer loops are shown in Table 16 below:

Table 16: EVPPI Results from both SAVI and Monte Carlo simulations with different numbers of iterations. AE: Adverse Events; MC: Monte Carlo; SAVI: Sheffield Accelerated Value of Information

	EVPPI per patient per year (€)		
Simulation	**AE Incidence**	**AE Costs**	**Resource Use Costs**
MC (100 inner/100 outer loops)	25.51	23.55	19.30
MC (200/100)	13.08	2.78	11.20
MC (300/100)	1.66	0.75	0.07
MC (100/200)	37.10	30.21	32.84
MC (100/300)	24.24	34.84	33.99
MC (100/400)	21.45	26.46	32.30
SAVI (3000 iterations)	754.25	532.91	66.75
SAVI (5000 iterations)	939.87	650.57	65.33

From Table 16, it is clear that the Monte Carlo method creates lower EVPPI estimates compared to the SAVI-results when assessing the current model. In addition, a comparison between MC results shows that the figures are not sufficiently stable at the relatively low numbers of inner and outer loops. For example, the relationship between the categories change between 300 and 400 outer loops (100 inner), where AE costs are associated with the highest EVPPI in one, and resource use costs in the other. Therefore, although the nonparametric regression approach of the SAVI-tool provides EVPPI estimates that are significantly higher than those generated by the MC approach, the SAVI-estimates may be considered more robust due to the higher number of samples that the results are based on (Strong et al, 2013).

6. Discussion

6.1 Main Findings

The aim of this thesis was to determine the cost-effectiveness of vortioxetine as compared with the current standards of care and placebo in first-line treatment of moderate to severe depression in Finland. The current standards of care were determined as escitalopram and venlafaxine. The base case deterministic results show that venlafaxine is expected to be the most cost-effective treatment alternative of those presented in this study. Venlafaxine treatment is associated with a lower cost and higher benefits than the chosen comparators. These differences, although relatively small, are driven by better efficacy and consequent mitigation of treatment-resistant cases as compared to the other treatments.

The main cost category is resource use including GP and psychiatry visits, as well as possible hospitalization and days spent in a ward. As resource use increases with prolonged treatment processes, the treatment associated with the shortest mean time of recovery is also the least costly and thus, cost-effective. This is why venlafaxine is considered cost-effective in the deterministic analysis, even though it is associated with higher incidence of adverse events, which is a main reason for preferring other treatments such as escitalopram in clinical practice (Duodecim, 2020).

Adverse events have a more significant effect in the scenario analysis considering a one-year time horizon, where venlafaxine is dominated by the other active treatment alternatives. Escitalopram was found to be the most affordable alternative in this scenario, while vortioxetine was associated with increased benefits at an increased cost (ICER: €44 616 against escitalopram). Adverse events are relatively more important in a shorter time perspective, due to the assumption that they occur during the first cycle after treatment initiation. As treatment switches occur primarily during the first year of the simulation, the effect of AEs is mitigated when considering longer time horizons.

Conclusions from the societal perspective were similar to the base case analysis, with venlafaxine dominating over a 20-year time horizon. Costs from a societal perspective were more than triple from the base case health care perspective and made up approximately 70% of overall costs related to depression. The substantial effect of productivity costs is in line with, for example, Ekman et al (2013), although their analysis from Sweden suggests an even higher proportion of costs arising from productivity losses. Nevertheless, the

results from this study underline the importance of societal costs in Finland and suggests these costs should always be included as scenarios in cost-effectiveness analyses regarding depression treatments.

The probabilistic sensitivity analysis illustrates the uncertainty associated with identifying the most cost-effective alternative. Although the deterministic results predict venlafaxine as the cost-effective option, the PSA reveals that it is only slightly over 50% likely to be cost-effective when compared with vortioxetine at a €20 000 threshold. In addition, the CEACs reveal that, when considered together, each active treatment alternative is associated with a similar probability between 30%-40% of being cost-effective. The impact of a threshold in decision-making is unclear in this case. On one hand, the relationship between treatments changes around the threshold of €25 000, as seen on the CEACs. On the other, none of the treatments emerge as definitively cost-(in)effective at any threshold, and never reach a probability of cost-effectiveness of over 37%. The results are characterized by a high level of uncertainty.

The value of information measures indicate that in order to remove uncertainty, the decision maker should be willing to spend at most €1343 per patient per year on new studies, given a €20 000 threshold. If all MDD patients are considered as the annual patient population (280 000 people, see introduction), this gives a total of more than €376 million per year to spend on additional studies. EVPPI measures give further direction to where studies may specifically be worthwhile: adverse event costs stand out as the category with the highest EVPPI of €940 per patient. This implies that a study including adverse event costs could potentially be conducted with an annual cost of approximately €263 million in order to reduce uncertainty of treatment decisions, according to the model. However, this value is to be interpreted with caution, as the relevant population may not be equal to the total number of patients on antidepressants (i.e. not all have severe depression, and not all are in acute treatment), and the value is heavily dependent on the modeling assumptions and structure, as demonstrated by the EVPPI values achieved using the Monte Carlo method. Nevertheless, the relatively high VOI-values indicate potential for future research in this area.

Despite the uncertainty displayed in the model, the results show that all antidepressant treatments under consideration are more effective and less costly than placebo in treating moderate to severe depression. This illustrates the importance of delivering appropriate and timely care to those in need, which is a specific challenge in depression treatment, due to its characteristic symptoms including apathy and lack of interest. Not

only does it improve the health and quality of life of the individual patient, but proper treatment may also relieve the burden on society caused by indirect costs, such as productivity losses.

6.2 Model Validity

There is a lack of studies available comparing the relevant treatments in a first-line treatment setting. Studies such as Soini et al (2017) focus on treatment decisions *after* the failure of SSRIs or other first-line therapies, which makes their results incomparable with the current study. Similar results to the current model have been obtained in the analysis by the Canadian Agency for Drugs and Technologies in Health (CADTH) in their reimbursement recommendation for vortioxetine (2020). They conclude, based on a one-year time horizon, that differences in treatment benefit between vortioxetine and other treatments (e.g. escitalopram) were minimal, and vortioxetine was substantially more expensive. The same result holds in the one-year scenario analysis of this study. Details of the model submitted to CADTH are not available, so further comparison is not possible for assessing cross validity.

This study uses methods seen in other cost-effectiveness models in the field of depression. For example, the implementation of a six-month maintenance phase has been in previous studies from both Finland (Soini et al, 2017) and elsewhere (Choi et al, 2016; NICE TA367, 2015), as well as 8-week cycles and the assumption of no relapse after achieving recovery (Choi et al, 2016; NICE TA367, 2015; Soini et al, 2017). Similarly, the assumption of AE occurrence only during the first cycle (Choi et al, 2016), and the omission of all-cause mortality are features seen in previous models. However, the current study has a longer time horizon than the references studies, which take a one-year time horizon, and has a novel way of accounting for disability pensions, specifically in Finland. Nevertheless, the lack of relevant comparisons limits possibilities of cross and external validation.

This study has taken the advice of the Evidence Review Group (ERG) commenting on Lundbeck's (vortioxetine patent holder) cost-effectiveness submission towards NICE TA367, which considered vortioxetine as a second-line treatment option. For example, the ERG critiqued the company for basing treatment change decisions solely on remission data, as clinicians base their decisions on response rather than remission. In addition, the ERG noted that half-cycling both costs and effects would be appropriate. These suggestions, among others were included in the current analysis as they were seen applicable to the first-line

treatment context as well. Considering second- and third-line treatments, the suggestion of ERG to use a proportional reduction of efficacy according to the STAR*D trial has been implemented in this thesis, as explained in Section 4.2.

The current model displays some of the key characteristics of treatment observed in practice. Due to the calibration of the recovery parameter from observational data, the table approximates real-world recovery rates in Finland after 11 years from the point of diagnosis, according to the study by Markkula et al (2015). The figures presented by the model are slightly more optimistic (i.e. more people have recovered), however this may be due to the assumptions of perfect patient adherence and lack of recurrence after recovery. On the other hand, the population in this model displays a severe form of depression, which predicts slightly lower recovery rates. Parameters regarding the most severe events in practice, suicide and early retirement, are also informed by observational registry data from Finland, which provides specific validity for these results in the Finnish context. A list comparing observational data with the predictions of the model is presented in Appendix 5. Although it was not possible to consult an impartial expert for assessing face validity in this study, the accordance with real-life observations might give some indication.

Internal validity was reviewed using the TECH-VER verification checklist (Buyukkaramikli et al, 2019) (see Appendix 4). This allowed for identifying potential sources of computational errors or other modeling inconsistencies. As a result of the process, one source of error was identified: a parameter had an inaccurate probability distribution. After correction, the model passed all the applicable tests of the TECH-VER verification checklist, as listed in Appendix 4.

6.3 Model Generalizability and Transferability

This model considers a specific population of moderate to severe MDD patients in Finland. As such, it may not be generalizable to a wider population of, for example, elders or adolescents. Patients younger than 18 and older than 64 have different treatment guidelines (Duodecim, 2020) including different dosages and preferences for certain treatments over others. There may also be differences in the procedure of initiating treatment switches, which are built into the model. Therefore, any attempts to generalize the findings of this model should be accompanied with the appropriate reflections on the differences between populations,

especially concerning resource use and, if taking a societal perspective, productivity costs. These same considerations should apply when addressing a population with a less severe form of MDD.

The model structure itself should be transferable to other countries, provided there are similar treatment guidelines as in Finland. The results of this study should, however, not be transferred directly to other contexts, as they rely heavily on data of, for example, resource use and costs of Finnish MDD patients specifically. Appropriate research should be conducted on the input parameters in a specific context, before attempting to interpret the current model in another setting.

6.4 Strengths of the Study

This analysis incorporates the most comprehensive literature review and network meta-analysis based on randomized controlled trials of antidepressants to date (Cipriani et al, 2018). Instead of relying on separate RCTs or indirect comparisons, using the results of Cipriani et al (2018) allows for consistency and robustness of key parameters, namely: response and remission rates, and adverse events parameters. Further, this study has used data specifically applicable to Finland, where appropriate. These parameters include utility values, suicide rates, and treatment-resistant recovery rates. The use of Finland-specific data provides further validation of the model's applicability in its intended context.

This study includes a cohort Markov model that can capture the long-term costs and effects of first-line treatment of major depression. Further, it includes a modifiable time horizon, which allows for examining the effect of extrapolating results into the future. Whereas past studies have typically included a one-year time horizon, the current model accounts for the recurring nature of more severe depression with a time horizon of up to 25 years. It also includes a value of information analysis, which, to the author's knowledge, has not been conducted in the context of antidepressants in Finland.

Finally, this study captures some of the suggestions by the ERG concerning the methodology of NICE TA367. Directly implementing the improvements presented by an established HTA body improves the structural quality of the model and provides credibility to the study. In addition, following and passing the verification process directed by the TECH-VER demonstrates strength in the structure of the model and reduces the likelihood of computational errors. The TECH-VER also promotes transparency, which increases the replicability of the study along with the methods and theory presented in this thesis.

6.5 Assumptions and Limitations

The most obvious limitation of this study is that it does not include all treatment alternatives available for first-line depression pharmacotherapy in Finland. Although the chosen comparators are able to represent the most common categories of antidepressants (i.e. SSRIs and SNRIs), they do not provide an exhaustive comparison of available treatment options. In addition, while relying on a single data source has its benefits (see above), the study is susceptible to any methodological inconsistencies that may be present in the study by Cipriani et al (2018). For example, it may be questioned whether the study population is truly representative to that of the Finnish patients discussed in this study. If not, the differences between the study population and the one represented in clinical practice should be assessed to estimate the potential bias introduced to the model presented in this thesis.

This study makes some assumptions and omissions that may influence the results and the conclusions drawn from them. Dose titration, for example, is common practice with antidepressants (Duodecim, 2020) and may in some cases affect treatment switch or continuation decisions. In order to retain a level of simplicity and transparency, dosage has been assumed constant throughout the model. Further, patients are assumed to display perfect treatment adherence. Treatment adherence is one of the most significant issues in treating depression (Duodecim, 2020), as patients may discontinue treatment from their own initiative without consulting the doctor. This assumption may favor treatment regimens with higher incidence of adverse events (venlafaxine).

Adverse events are assumed to have an additive effect on health-related quality-of-life utilities in the model. This assumption may be questioned in practice, as the additional burden experienced from concurrent adverse events may not equally severe as the burden experienced due to the initial event. For example, when assessed separately, headache and fatigue are both associated with a disutility of 0.08 QALYs. When considered additively, an individual suffering from both would experience a disutility of 0.16. However, if the events were to be assessed together, the combination of headache and fatigue might arguably be less or more than 0.16. However, this assumption is necessary due to the challenges involved with modeling all combination of adverse events, including a lack of relevant data.

This study assumes that patients have a primary diagnosis of depression, and do not suffer from other psychiatric disorders, such as anxiety or psychosis. In practice, however, it is common for patients to display mental comorbidities and especially the combination of depression and anxiety (Markkula et al, 2015). The

effect of secondary diagnoses is not examined in this paper, nor are any possible differences that might arise in treatment decisions. This assumption has been made due to both data validity, as well as comparability with previous studies.

All-cause mortality has been omitted due to presumed insignificance for the analysis and conclusions drawn in this thesis. In addition, this model has omitted recurrences after recovery, which may have a significant effect on results, especially considering the relatively long time horizon of the base case analysis. However, modeling recurrences would have required additional assumptions, both with regards to input parameters and model structure. Further, the effect of recurrences is expected to be similar across all treatment alternatives.

The number or non-fatal suicide attempts cannot be conclusively estimated in Finland, as only hospitalized cases are included in the appropriate register. Nevertheless, the World Health Organization has estimated that for each fatal suicide, there are approximately 20 non-fatal suicide attempts (Solin et al, 2019). In addition to difficulty in estimating the occurrence, it is challenging to estimate associated costs with suicide attempts, regardless of their outcome. Therefore, non-fatal suicides have been omitted from the analysis, and the costs associated with fatal events have not been included. A Finnish study estimated that suicide is associated with at least 11 different cost categories, and studies from Finland and elsewhere display large variance in their estimation of total costs from €60 000 up to €1.3 million (Solin et al, 2019). This apparent uncertainty is the main motivation for excluding these costs from the current analysis, although suicide mortality has been included due to the perceived importance of considering the most dramatic outcome for an individual.

Value of information measures may be influenced by the assumption of independence between input variables. Naversnik and Rojnik (2012) conclude that, while assuming independence between parameters (i.e. not accounting for correlation) rarely alters the optimal decision estimated by a model, correlated variables may cause VOI-measures to be highly under- or overestimated, depending on the strength and direction of correlation. In the current study, several sources of correlation can be identified, including the negative correlation between costs and effects, resulting from an increased use of resources in health states associated with lower utilities. For adverse events, the incidence and costs associated with one treatment influence the

other treatment alternatives as well, due to the treatment algorithm in subsequent lines of treatment. Nevertheless, statistical analyses related to correlation are beyond the scope of this thesis.

Finally, in the probabilistic sensitivity analysis, some parameters were assumed an uncertainty of 20% of the mean value. While it has been necessary to make assumptions where data is not available, this value is essentially arbitrary and has a significant effect on the probabilistic results. Most notably, it affects the value of information measures. Appropriate caution should be exercised when interpreting these results.

6.6 Future Research

The results of this research imply that current guidelines in Finland are justified with regards to withholding from making any definite recommendations of one antidepressant over another. Nevertheless, future research in the area of cost-effectiveness could be worthwhile, as indicated by the relatively high EVPI figures estimated by the analysis. This study has demonstrated that the use of vortioxetine should be considered as an equivalent option to escitalopram and venlafaxine in the first-line treatment of major depressive disorder. More research needs to be done in order to identify the patient population that stands to gain the most from each individual treatment, as they all have different mechanisms of action.

There is a clear need for cost-effectiveness analyses in the field of depression in Finland. Although the condition is highly relevant from a public health perspective, existing studies are scarce. In addition to economic analyses, future research may be well placed in the incidence and associated costs of adverse events, following their relatively high EVPPI figures. For example, an observational study of adverse events incidence in clinical practice, and their associated costs, could provide accurate real-world evidence, and reduce related uncertainty. Both economic and clinical studies in different contexts, with different comparators and with different patient populations could make future treatment guidelines more specific and ease the identification of the most suitable treatment for each patient.

This study considers three different antidepressants and presents a placebo comparison. The analysis presented illustrates the difference between patients receiving appropriate treatment and those who do not. The simulated placebo cohort incurs more costs and experiences less benefits than each of the active treatment alternatives. In practice, those not seeking or receiving help might not expect even those results presented under placebo in this study, as they will not experience a placebo effect. In addition, the data informing

placebo-related parameters in this thesis come from patients who have been part of clinical studies, and therefore likely to have a care contact. Thus, it is reasonable to assume that patients not receiving care are worse off with their symptoms, highlighting the necessity to receive appropriate care. Although it is not the primary focus of this study, it has illustrated the need for future research to ensure depressed patients receive appropriate and timely care, as lack of it may be costly both on an individual and on a societal level.

7. Conclusion

This study has illustrated that in the treatment of moderate to severe major depressive disorder of adult patients in Finland, vortioxetine is more likely to be cost-effective than escitalopram, venlafaxine, or placebo at cost-effectiveness thresholds higher than €30 000. The study also reveals that the choice of antidepressant is associated with a high level of uncertainty regarding both costs and effects, and thus cost-effectiveness. In addition, it is illustrated that societal costs account for approximately 70% of costs incurred by depression. The significance of societal costs demonstrates the public health relevance of major depressive disorder in Finland and urges the need for further cost-effectiveness research in an area were existing literature is scarce.

References

Attema, A., Brouwer, W., Claxton, K. (2018). *Discounting in Economic Evaluations*. PharmacoEconomics (2018) 36:745-758. Doi: 10.007/s40273-018-0672-z

Arrow, K., Lind, R. (1970). *Uncertainty and the Evaluation of Public Investment Decisions*, Amer. Econ. Rev., June 1970, 60, 364-78.

Baldwin, D. S., Reines, E. H., Guiton, C., & Weiller, E. (2007). *Escitalopram Therapy for Major Depression and Anxiety Disorders.* Annals of Pharmacotherapy, 41(10), 1583–1592. doi.org/10.1345/aph.1K089

Baldwin, D., Chrones, L., Florea, I., Nielsen, R., Nomikos, G., Palo, W., Reines, E. (2016). *The safety and tolerability of vortioxetine: Analysis of data from randomized placebo-controlled trials and open-label extension studies.* Journal of Psychopharmacology 2016, 30(3) 242-252. Doi: 10.1177/026881116628440

Bromet, E., Andrade, L. H., Hwang, I., … Kessler, R. C. (2011). *Cross-national epidemiology of DSM-IV major depressive episode*. BMC Med. 2011 Jul 26;9:90 doi: 10.1186/1741-7015-9-90

Boulenger, J-P., Loft, H., Florea, I. (2012). *A randomized clinical study of Lu AA21004 in the prevention of relapse in patients with major depressive disorder.* Journal of Psychopharmacology 26(11) 1408-1416. Doi: 10.1177/0269881112441866

Büyükkaramikli N., Rutten-van Mölken M., Severens H., Al M. (2019).: *TECH-VER: A verification checklist to reduce errors in models and improve their credibility*. Pharmacoeconomics. 37(11):1391–1408. https ://doi.org/10.1007/s4027 3-019-00844 -y

Choi, S-E., Brignone, M., Seong. C., Rangrhee, J., Campbell, R., Francois, C., Milea, D. (2016) *Cost-effectiveness of vortioxetine versus venlafaxine (extended release) in the treatment of major depressive disorder in South Korea*, Expert Review of Pharmacoeconomics & Outcomes Research, 16:5, 629-638, DOI:10.1586/14737167.2016.1128830

Cipriani A., La F., Furukawa T., … Barbui, C. (2010). *Sertraline versus other antidepressive agents for depression.* Cochrane Database Syst Rev. 2010;1:CD006117.

Cipriani, A., Furukawa T., Salanti, G., ... Geddes, J. (2018). *Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis.* Lancet 2018; 391:1357-66. doi: 10.1016/S0140-6736(17)32802-7

Citrome L. (2014). *Vortioxetine for major depressive disorder: a systematic review of the efficacy and safety profile for this newly approved antidepressant – what is the number needed to treat, number needed to harm and likelihood to be helped or harmed?* Int J Clin Pract 2014; 68(1): 60–82. doi: 10.1111/ijcp.12350

Cleemput, I., Neyt, M., Thiry, N., De Laet, C., Leys, M. (2011). *Using threshold values for cost per quality-adjusted life-year gained in healthcare decisions.* Int J Tech Ass in Health Care, 271 (2011), 71-

76. doi:10.1017/S0266462310001194

Cronquist-Christensen, M., Munro, V. (2018). *Cost per successfully treated patient for vortioxetine versus duloxetine in adults with major depressive disorder: an analysis of the complete symptoms of depression and functional outcome*, Current Medical Research and Opinion, 34:4, 593-600, DOI: 10.1080/03007995.2017.1416952

Deeks J, Higgins J, Altman D. (editors). Chapter 10: *Analysing data and undertaking meta-analyses.* In: Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0. Cochrane, 2019. Retrieved May 5, 2020 from training.cochrane.org/handbook

Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). *Methods for the Economic Evaluation of Health Care Programmes*. Oxford University Press.

Duodecim. (2020). *Depression*. *Current Care Guidelines*. Working group set up by the Finnish Medical Society Duodecim and the Finnish Cardiac Society. Helsinki: The Finnish Medical Society Duodecim. Retrieved January 8, 2020 from kaypahoito.fi

Eddy, D., Hollingworth, W., Caro, J., Tsevat, J., McDonald, K., Wong, J., & ISPOR-SMDM Modeling Good Research Practices Task Force (2012). *Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7.* Medical decision making: an international journal of the Society for Medical Decision Making, 32(5), 733–743. https://doi-org.eur.idm.oclc.org/10.1177/0272989X12454579

Ekman, M., Granström, O., Omerov, S., Jacob, J., Landen, M. (2013). *The societal cost of depression: Evidence from 10,000 Swedish patients in psychiatric care.* Journal of Affective Disorders (2015), 150, 790-797: doi: 10.106/j.jad.2013.03.003

Ferrari, A. J., Somerville, A., Baxter, A. J., Norman, R. E., Patten, S. B., Vos, T. & Whiteford, H. A. (2013). *Global variation in the prevalence and incidence of major depressive disorder: a systematic review of the epidemiological literature*. Psychological Medicine (2013), 43, 471-481. Doi: 10.1017/S0033291712001511

Finnish Centre for Pensions. (2020). *Earnings-related pension recipients in Finland.* Finnish Centre for Pensions. Retrieved May 9, 2020 from etk.fi/en

Finnish Medicines Agency (Fimea) (2012). *Fimean suositus lääkkeiden hoidollisen ja taloudellisen arvon arvioinnista [Fimea's recommendation for the medical and economic appraisal of medicines]*. Fimea. Retrieved February 4, 2020 from julkari.fi

Finnish Medicines Agency (Fimea) & The Social Insurance Institution (KELA). (2019). *Finnish Statistics on Medicines 2018.* Fimea. Retrieved March 26, 2020 from julkari.fi

Finnish Medicines Agency (Fimea) (2020). *Myyntiluvat [Market Authorization].* Fimea. Retrieved March 30, 2020 from fimea.fi

Furukawa, T., Cipriani, A., Atkinson, L., Leucht, S., … Salanti, G. (2016). *Placebo response rates in antidepressant trials: a systematic review of published and unpublished double-blind randomized controlled studies.* Lancet Psychiatry 2016; 3:1059-66. doi: 10.1016/S2215-0366(16)30307-8

Heiskanen, T., Huttunen, M. O., & Tuulari, J. (2011). *Masennus [Depression]*. Helsinki: Duodecim.

Hoaglin, D., Hawkins, N., Jansen, J., Scott, D., Itzler, R., Cappelleri, J.., … Barrett, A. (2011). *Conducting Indirect-Treatment-Comparison and Network-Meta-Analysis Studies: Report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: Part 2*. Value in Health, 14(4), 429–437. doi: 10.1016/j.jval.2011.01.011

Huttunen, M. (2018). *Masennustilat eli depressiot [States of depression].* Duodecim Terveyskirjasto. Retrieved August 23, 2019 from terveyskirjasto.fi

Ikäheimo, P. (2014). *Vortioksetiini [Vortioxetine].* SIC! Lääketietoa Fimeasta 2/14. Fimea. Retrieved May 2nd from sic.fimea.fi

Jakobsen, J.C., Katakam, K.K., Schou, A., Hellmuth, S.G., … Gluud, C. (2017). *Selective serotonin reuptake inhibitors versus placebo in patients with major depressive disorder. A systematic review with meta-analysis and Trial Sequential Analysis.* BMC Psychiatry 2017: 17:58 doi: 10.1186/s12888-016-1173-2

Kansaneläkelaitos (KELA). [Institute of Social Welfare]. (2020). *Lääkehaku [Drug search].* KELA. Retrieved March 31, 2020 from asiointi.kela.fi

Karlsson, H. (2012). *Masennus systeemisairautena [Depression as a systemic disease].* The Finnish Medical Society Duodecim. Duodecim Medical Journal 2012: 128(6): 622-6

Kuehner, C. (2017). *Why is depression more common among women than among men?* The Lancet Psychiatry, Volume 4, Issue 2, 2017, Pages 146-158 doi: 10.1016/S2215-0366(16)30263-2

Li, T., Puhan, M.A., Vedula, S.S. et al. *(2011) Network meta-analysis-highly attractive but more methodological research is needed*. BMC Med 9, 79 (2011). doi: 10.1186/1741-7015-9-79

Lääkkeiden hintalautakunta (HILA) (2020). *Usein kysytyt kysymykset. [Frequently Asked Questions].* HILA. Retrieved March 31, 2020 from hila.fi

Markkula, N., Suvisaari, J., Saarni, S.I., et al. (2015). *Prevalence and correlates of major depressive disorder and dysthymia in an eleven-year follow-up – results from the Finnish Health 2011 Survey*. J Affect Disord 2015;173:73–80

Markkula, N. (2016). *Prevalence, Predictors and Prognosis of Depressive Disorders in the General Population.* Academic Dissertation. University of Helsinki. Retrieved March 2, 2020 from http://urn.fi/URN:ISBN:978-951-51-2071-7

Mavridis, D., Giannatsi, M., Cipriani, A., Salanti, G. (2015). *A primer on network meta-analysis with emphasis on mental health.* Evid Based Mental Health 2015;18:40-46.

McIntyre, R. (2017). *The role of new antidepressants in clinical practice in Canada: a brief review of vortioxetine, levomilnacipran ER, and vilazodone.* Neuropsychiatr Dis Treat. 2017; 13: 2913–2919. doi: 10.2147/NDT.S150589

National Collaborating Centre for Mental Health (NCCMH). (2010). *Depression: the treatment and management of depression in adults (Updated Edition)*. Leicester (UK): The British Psychological Society & The Royal College of Psychiatrists; 2010.

National Institute of Care Excellence (NICE). (2018). *NICE TA367, Vortioxetine for treating major depressive episodes.* NICE. Retrieved May 10 from nice.org.uk

Naversnik, K., Rojnik, K. (2012). *Handling Input Correlations in Pharmacoeconomic Models.* Value in Health., 15(3). p. 540-549. doi: 10.1016/j.jval.2011.12.008

Osby, U., Brandt, L., Correia, N., Ekbom, A., & Sparén, P. (2001). *Excess mortality in bipolar and unipolar disorder in Sweden*. Archives of general psychiatry, 58(9), 844–850. doi: 10.1001/archpsyc.58.9.844

Rapaport, M., Bose, A., Zheng, H. (2004). *Escitalopram Continuation Treatment Prevents Relapse of Depressive Episodes.* J Clin Psychiatry 2004;65: 44-49

Riihimäki, K.A., Vuorilehto, M.S, Melartin, T.K., Isometsä, E.T. (2014). *Five-year outcome of major depressive disorder in primary health care.* Psychological Medicine (2014), 44, 1369-1379. doi: 10.1017/S0033291711002303

Saarni SI, Harkanen T, Sintonen H, et al. (2006). *The impact of 29 chronic conditions on health-related quality of life: a general population survey in Finland using 15D and EQ-5D*. Qual Life Res. 2006;15 (8):1403–1414.

Simon, J., Aguiar, L., Kunz, N., Lei, D. (2004). *Extended-release venlafaxine in relapse prevention for patients with major depressive disorder.* Journal of Psychiatric Research 38 (2004) 249-257

Soini, E., Hallinen, T., Brignone, M., Campbell, R., Diamand, F., Cure, S., Aalto-Setälä, M., Danchenko, N., Koponen, H., Kolasa, K., (2017*). Cost-utility analysis of vortioxetine versus agomelatine, bupropion SR, sertraline and venlafaxine XR after treatment switch in major depressive disorder in Finland.* Expert Review of Pharmacoeconomics & Outcomes Research, 17:3, 293-302, DOI: 10.1080/14737167.2017.1240617

Solin, P., Jokinen, J., Partonen, T., Reinikainen, J., Seppänen, A., Tamminen, N. (2019). *Itsemurhien ehkäisy: koulutus perusterveydenhuollon ammattilaisille. [Suicide prevention: training for primary care professionals].* The Institute of Health and Welfare (THL). ISBN 978-952-343-332-8.

Sosiaali- ja terveysministeriö (STM) [Ministry of Social Affairs and Health]. (2013b). *Terveydenhuolto Suomessa [Healthcare in Finland].* STM. Retrieved March 27, 2020 from julkari.fi

Sosiaali- ja terveysministeriö (STM) [Ministry of Social Affairs and Health]. (2020). *Terveydenhuollon Maksut & Lääkekorvaukset* [Health Care Payments & Medical Reimbursements]. STM. Retrieved March 27, 2020 from stm.fi

Sotelo, J.L., Nemeroff, C.B. (2017). *Depression as a systemic disease.* Personalized Medicine in Psychiatry 1-2; 2017; 11-25. doi: 10.1016/j.pmip.2016.11.002

Strong M, Oakley JE, Brennan A. (2014). *Estimating multi-parameter partial Expected Value of Perfect Information from a probabilistic sensitivity analysis sample: a non-parametric regression approach*. Medical Decision Making. 2014;34(3):311-26. Available from savi.shef.ac.uk

Terveyden ja Hyvinvoinnin Laitos (THL) [Institute of Health and Welfare]. (2012). *Psykiatrian luokituskäsikirja [The Handbook of Psychiatric Classifications]*. THL. Retrieved January 7, 2020 from julkari.fi

Terveyden ja Hyvinvoinnin Laitos (THL) [Institute of Health and Welfare]. (2020). *Depressiolääkkeistä korvausta saaneet 18-64 / 1000 vastaavanikäistä* [Patients receiving reimbursements for antidepressants, 18-64-year-olds / per 1000], SOTKAnet Statistics and Indicator Bank. Retrieved March 28, 2020 from sotkanet.fi

Tuffaha, H., Strong, M., Gordon, L., Scuffham, P. (2016). *Efficient Value of Information Calculation Using a Nonparametric Regression Approach: An Applied Perspective.* Value in Health 19(2016): 505-509. Doi: http://dx.doi.org/10.1016/j.jval.2016.01.011

Zhang J, Yu KF. (1998). *What's the Relative Risk? A Method of Correcting the Odds Ratio in Cohort Studies of Common Outcomes*. JAMA. 1998;280(19):1690–1691. doi:10.1001/jama.280.19.1690

Appendix 1: Systematic Literature Review Search

1. PubMed

Query #1: ("vortioxetine"[MeSH Terms] OR "vortioxetine"[All Fields]) AND (("depressive disorder"[MeSH Terms] OR ("depressive"[All Fields] AND "disorder"[All Fields]) OR "depressive disorder"[All Fields] OR "depression"[All Fields] OR "depression"[MeSH Terms]) OR MDD[All Fields]) AND Clinical Trial[ptyp]

Query #2: "escitalopram"[Title] AND ("Major depressive disorder"[Title/Abstract] OR depression[Title/Abstract]) AND randomized[title/abstract] AND Clinical Trial[ptyp]

Query #3: "venlafaxine"[Title] AND ("major depressive disorder"[Title/Abstract] OR depression[Title/Abstract]) AND Clinical Trial[ptyp]

2. Embase

Query #1: (vortioxetine:ti OR 'escitalopram':ti OR venlafaxine:ti) AND ('major depressive disorder':ab,ti OR depression:ab,ti) AND [randomized controlled trial]/lim

Query #2: #1 AND [adult]/lim

Query #3: #1 AND [adult]/lim AND [embase]/lim NOT

3. ClinicalTrials.gov

Query #1: vortioxetine OR escitalopram OR venlafaxine, filters: Adults, With Results

4. Wiley Online Library

Query #1: "vortioxetine OR escitalopram OR venlafaxine" in Title and ""major depressive disorder" OR depression" in Keywords

Appendix 2: Literature Review Results

Study; drug, mg/day	*Mean age (years)*	*Sex, % female*	*Duration, weeks*	*n*	*Baseline MADRS*	*Study location*	*MADRS remission, %*	*Entry score*
Alvarez et al			6			Asia/Australia/ Europe		30
PBO	42.0	65.7		105	33.9 +/- 2.7		27	
VTX, 5	43.8	64.8		108	34.1 +/- 2.6		49	
VTX, 10	42.3	66.0		100	34.0 +/- 2.8		49	
VFX, 225	45.0	54.9		113	34.2 +/- 3.1		55	
Baldwin et al			8			Asia/Europe		26
PBO	43.4	69.6		148	31.7 +/- 4.3			
VTX, 5	44.7	66.2		157	32.7 +/- 4.8		36	
VTX, 10	45.2	68.1		151	31.8 +/- 3.9		36	
Boulenger et al			8			Europe		26
PBO	48.1	69.6		158	31.5 +/- 3.6		19	
VTX, 15	47.0	64.2		151	31.8 +/- 3.4		34.9	
VTX, 20	46.2	60.3		151	31.2 +/- 3.4		38.4	
Colonna et al	46	73	8/24			Europe		22/30*
*ESC, 10, 8 wk**	n/a	n/a		80	33.5		33	
*ESC, 10, 24 wk**	n/a	n/a		80	33.5		70	
Hewett et al			8			Europe/USA		18†
PBO	41.8	72		197	30.4 (0.34)		32	
VFX, 75-150	42.7	68		187	30.0 (0.34)		51	
Hewett et al (2010)			8			USA		18†
PBO	44.5	67		187	30.6 (0.38)		38	
VFX, 75-150	44.1	68		198	30.1 (0.37)		56	
Jacobsen et al			8			USA		26
PBO	42.3	70.1		157	32.0 +/- 4.0		14.2	
VTX, 10	43.1	76.1		155	32.3 +/- 4.5		21.4	
VTX, 20	43.1	71.3		150	32.4 +/- 4.3		22.3	
Jain et al			6			USA		≥30
PBO	42.4	54.7		300	32.2		32.2	

					+/- 5.5			
VTX, 5	42.5	62.0		300	32.7 +/- 5.4		29.1	
Montgomery et al			8			Europe		18
ESC, 10-20	49	73		146	28.7 +/- 5.0		77.4	
VFX, 75-150	47	71		142	29.0 +/- 5.4		79.6	
Moore et al			8			Europe		30
ESC, 20	44.1	71.7		138	36.3 +/- 4.8		56.1	
NCT00735709			8			Global		26
PBO	46.4	61.4		140	30.6 (2.89)		16.5	
VTX, 5	47.3	62.1		140	30.6 (2.83)		28.8	
VTX,10	46.4	60.7		140	31.6 (3.83)		26.6	
NCT02389816			8			Japan		26
PBO	39.5	43.9		164	30.5 (3.87)		21.1	
VTX, 10	40.0	43.6		165	30.8 (3.73)		32.1	
VTX, 20	40.4	48.8		164	30.6 (3.62)		30.9	
Nishimura et al			8			Asia/Europe		26
PBO	43.6	59.9		152	31.6 +/- 3.56		26.7	
VTX, 5	44.2	68.1		144	31.6 +/- 3.67		24.6	
VTX, 10	45.7	62.0		150	31.8 +/- 4.02		29.3	
VTX, 20	44.0	60.4		154	31.7 +/- 3.73		30.9	
Ventura et al			8			USA		22
ESC, 10	40.6	54.8		104	29.5 +/- 0.4		60	
Yevtushenko et al			6			Russia		25
ESC, 10	35.2	61.1		108	34.78		82.4	

ESC: Escitalopram; PBO: Placebo; VFX: Venlafaxine; VTX: Vortioxetine

**Severely depressed (MADRS 30) patients a specified separate subgroup in the study, results from that subgroup*

† Assessed on HAM-D scale

Appendix 3: List of Parameters

List of all parameters directly informed by the literature. Underlined standard errors are according to the assumption of 20% from the mean value.

Parameter	Deterministic value	OR (RR)	Distribution	SE (SD)	Description	Source
Response rates						
Vortioxetine	0.48	1.66 (1.34)	Lognormal	0.07	Probability of achieving at least a 50% reduction in HAMD/MADRS score	Cipriani et al (2018)
Escitalopram	0.49	1.68 (1.35)	Lognormal	0.06		
Venlafaxine	0.50	1.78 (1.39)	Lognormal	0.05		
Placebo	0.36		Beta	0.02		Furukawa, Cipriani et al (2016)
Remission rates						
Vortioxetine	0.34	1.49 (1.32)	Lognormal	0.07	Probability of achieving a score of 12 (MADRS) or 7 (HAMD) or less	Cipriani et al (2018)
Escitalopram	0.36	1.64 (1.41)	Lognormal	0.06		
Venlafaxine	0.37	1.70 (1.44)	Lognormal	0.05		
Placebo	0.25		Beta	(0.01)		Jakobsen et al (2017)
Remission during recurrent depression	0.10		Beta	(0.02)		Rush et al (2006)
Relapse rates						
Vortioxetine	0.09	(0.50)	Lognormal	0.24	Probability of experiencing a recurrence during the respective depression period	Boulenger et al (2012)
Escitalopram	0.10	(0.56)	Lognormal	0.04		Rapaport et al (2004)
Venlafaxine	0.09	(0.50)	Lognormal	0.03		Simon et al (2004)
Placebo	0.17		Beta	0.05		Boulenger et al (2012)
2nd line	0.14		Beta	0.03		Soini et al (2017), TA367
Recurrent	0.14		Beta	0.03		Same as above
Recovery rate, **recurrent depression**	0.30		Beta	0.06	Probability of recovering from treatment resistant depression (per cycle)	Markkula et al (2011)
Utilities						
Depression	0.51		Beta	0.10		Soini et al (2017)
Remission	0.84		Beta	0.17		

Sustained depression (2nd line)	0.51		Beta	0.10	HRQoL utility value associated with the respective health state	
Recurrent depression (3rd+ line)	0.51		Beta	0.10		Assumed equal to baseline
Drug acquisition costs (€)						
Vortioxetine	63.85		Finnish law requires the same price by companies to all pharmacies in Finland. Hence, no uncertainty.		Cost of acquiring the respective drugs (per cycle)	Social insurance institute (KELA) (2020)
Escitalopram	2.51					
Venlafaxine	17.63					
AE Dropouts						
Vortioxetine	0.06	1.64 (1.60)	Lognormal	0.14	Probability of dropping out of (i.e. switching) treatment due to AEs	Cipriani et al (2018)
Escitalopram	0.06	1.72 (1.68)	Lognormal	0.11		
Venlafaxine	0.10	2.95 (2.76)	Lognormal	0.09		
Placebo	0.04		Beta	(0.004)		Baldwin et al (2016)
AE Costs (per cycle) (€, 2011)						
Constipation	13.50		Gamma	2.70	Cost associated with managing each adverse event	Unit costs of care from Kapiainen et al (2014). Resource use assumed by author.
Diarrhea	13.50		Gamma	2.70		
Dizziness	27.00		Gamma	5.40		
Dry mouth	13.50		Gamma	2.70		
Fatigue	13.50		Gamma	2.70		
Headache	13.50		Gamma	2.70		
Hyperhidrosis	27.00		Gamma	5.40		
Insomnia	27.00		Gamma	5.40		
Nausea	13.50		Gamma	2.70		
Sexual dysfunction	127.00		Gamma	25.40		
Somnolence	13.50		Gamma	2.70		
Nasopharyngitis	13.50		Gamma	2.70		
AE Incidence (vortioxetine)						
Constipation	0.04		Beta	(0.005)	Probability of suffering from each respective treatment-related AE during vortioxetine treatment	Baldwin et al (2016)
Diarrhea	0.07		Beta	(0.006)		
Dizziness	0.06		Beta	(0.005)		
Dry mouth	0.06		Beta	(0.005)		
Fatigue	0.03		Beta	(0.004)		
Headache	0.13		Beta	(0.008)		
Hyperhidrosis	0.02		Beta	(0.003)		
Insomnia	0.03		Beta	(0.004)		
Nausea	0.27		Beta	(0.010)		
Sexual dysfunction	0.02		Beta	(0.003)		
Somnolence	0.03		Beta	0.004)		
Nasopharyngitis	No significant incidence of nasopharyngitis reported					

AE Incidence (escitalopram)						
Constipation	No significant incidence of constipation reported					
Diarrhea	0.08		Beta	(0.008)	Probability of suffering from each respective treatment-related AE during escitalopram treatment	Baldwin et al (2007)
Dizziness	0.08		Beta	(0.009)		
Dry mouth	0.07		Beta	(0.008)		
Fatigue	0.10		Beta	(0.009)		
Headache	0.19		Beta	(0.012)		
Hyperhidrosis	0.06		Beta	(0.008)		
Insomnia	0.08		Beta	(0.008)		
Nausea	0.21		Beta	(0.012)		
Sexual dysfunction	0.02		Beta	(0.004)		
Somnolence	0.06		Beta	(0.007)		
Nasopharyngitis	0.06		Beta	(0.007)		
AE Incidence (venlafaxine)						
Constipation	0.10		Beta	0.028	Probability of suffering from each respective treatment-related AE during venlafaxine treatment	Baldwin et al (2007, 2016)
Diarrhea	0.06		Beta	0.012		
Dizziness	0.08		Beta	0.014		
Dry mouth	0.12		Beta	0.017		
Fatigue	0.08		Beta	0.014		
Headache	0.17		Beta	0.020		
Hyperhidrosis	0.15		Beta	0.033		
Insomnia	0.11		Beta	0.017		
Nausea	0.29		Beta	0.024		
Sexual dysfunction	0.10		Beta	0.016		
Somnolence	0.03		Beta	0.010		
Nasopharyngitis	No significant incidence of nasopharyngitis reported					
AE Incidence (placebo)						
Constipation	0.03		Beta	(0.004)	Probability of suffering from each respective treatment-related AE during placebo treatment	Baldwin et al (2007, 2016)
Diarrhea	0.05		Beta	(0.005)		
Dizziness	0.06		Beta	(0.005)		
Dry mouth	0.06		Beta	(0.006)		
Fatigue	0.03		Beta	(0.004)		
Headache	0.13		Beta	(0.008)		
Hyperhidrosis	0.02		Beta	(0.003)		
Insomnia	0.04		Beta	(0.005)		
Nausea	0.08		Beta	(0.006)		
Sexual dysfunction	0.01		Beta	(0.002)		
Somnolence	0.02		Beta	(0.004)		
Nasopharyngitis	0.03		Beta	(0.007)		
AE disutilities						
Constipation	0.10		Beta	0.020	Disutility associated with respective adverse event	Soini et al (2017)
Diarrhea	0.10		Beta	0.020		
Dizziness	0.10		Beta	0.020		
Dry mouth	0.10		Beta	0.020		

Fatigue	0.08		Beta	0.016		
Headache	0.08		Beta	0.016		
Hyperhidrosis	0.10		Beta	0.020		
Insomnia	0.08		Beta	0.016		
Nausea	0.10		Beta	0.020		
Sexual dysfunction	0.13		Beta	0.026		
Somnolence	0.12		Beta	0.024		
Nasopharyngitis	0.08		Beta	0.016		
Resource Use Costs (€, 2011)						
GP visits	83.00		Gamma	21.01	Cost associated with respective health care services	Kapiainen et al (2014)
Psychiatrist visits	162.00		Gamma	41.00		
Psychotherapy or counseling	186.00		Gamma	32.00		
Psychiatric ward (per day)	408.00		Gamma	47.00		
Outpatient hospitalization	234.00		Gamma	26.96		
Absenteeism (day)	168.00		Gamma	33.60		
Resource use (per cycle) (remission)						
GP visits	0.35		Gamma	0.07	Assumed average resource use per respective health state	Soini et al (2017)
Psychiatrist visits	-0.15		Normal	0.03		
Psychotherapy or counseling	0.06		Gamma	0.01		
Psychiatric ward (per day)	0.90		Gamma	0.18		
Outpatient hospitalization	0.12		Gamma	0.02		
Absenteeism (day)	0.94		Gamma	0.19		
Resource use (per cycle) (relapse)						
GP visits	0.31		Gamma	0.06	Assumed average resource use per respective health state	Soini et al (2017)
Psychiatrist visits	0.39		Gamma	0.08		
Psychotherapy or counseling	0.45		Gamma	0.09		
Psychiatric ward (per day)	0.51		Gamma	0.10		
Outpatient hospitalization	0.07		Gamma	0.01		
Absenteeism (day)	13.49		Gamma	2.70		
Resource use (per cycle) (depressed)						
GP visits	4.88		Gamma	0.98	Assumed average resource use per respective health state	Soini et al (2017)
Psychiatrist visits	5.88		Gamma	1.18		
Psychotherapy or counseling	3.00		Gamma	0.60		
Psychiatric ward (per day)	8.00		Gamma	1.60		

Outpatient hospitalization	-		Gamma	-		
Absenteeism (day)	43.00		Gamma	8.60		
% Using resource while depressed						
GP visits	0.95		Beta	0.19	Assumed % of patients using respective health services in the depressive health state	Soini et al (2017)
Psychiatrist visits	0.10		Beta	0.02		
Psychotherapy or counseling	0.25		Beta	0.05		
Psychiatric ward (per day)	0.01		Beta	0.00…		
Outpatient hospitalization	-		Beta	-		
Absenteeism (day)	0.27		Beta	0.05		
Suicide						
Probability per cycle (remission)	0.000001		Beta	0.00	Probability of suicide in the respective health state (per cycle)	SotkaNET, Suicide mortality in Finland, 20-64 y/o
Probability per cycle (depressed)	0.00003		Beta	0.00…		

Appendix 4: Scenario Analyses

Table A3.1. *Societal Perspective, 20-year time horizon*

Undiscounted					
Treatment	**Cost (€)**	**QALYs**	**ICER (€) compared to**		
			Lowest Cost	**Next lowest cost**	**Relevant alternative**
Venlafaxine	94 141	18.151	-	-	-
Escitalopram	94 652	18.143	SD	SD	SD
Vortioxetine	94 834	18.145	SD	93 060	SD
Placebo	107 677	17.897	SD	SD	SD
Discounted (3%)					
	Cost (€)	**QALYs**	**ICER (€) compared to**		
			Lowest Cost	**Next lowest cost**	**Relevant alternative**
Venlafaxine	78 986	13.518	-	-	-
Escitalopram	79 408	13.511	SD	SD	SD
Vortioxetine	79 583	13.513	SD	85 877	SD
Placebo	90 215	13.303	SD	SD	SD

Table A3.2. *1-, 5-, 10-, 15-, 25- year time horizon*

1-year, undiscounted					
Treatment	**Cost (€)**	**QALYs**	**ICER (€) compared to**		
			Lowest Cost	**Next lowest cost**	**Relevant alternative**
Escitalopram	5 097	0.701	-	-	-
Venlafaxine	5 105	0.700	SD	SD	SD
Vortioxetine	5 204	0.703	44 616	28 620	44 616€
Placebo	5 437	0.676	SD	SD	SD
5 years, discounted (3%)					
Treatment	**Cost (€)**	**QALYs**	**ICER (€) compared to**		
			Lowest Cost	**Next lowest cost**	**Relevant alternative**
Venlafaxine	15 425	3.698	-	-	-
Escitalopram	15 483	3.695	SD	SD	SD
Vortioxetine	15 592	3.698	SD	49 225	SD
Placebo	16 984	3.569	SD	SD	SD
10 years, discounted (3%)					
Treatment	**Cost (€)**	**QALYs**	**ICER (€) compared to**		
			Lowest Cost	**Next lowest cost**	**Relevant alternative**
Venlafaxine	21 085	7.287	-	-	-
Escitalopram	21 179	7.281	SD	SD	SD
Vortioxetine	21 289	7.284	SD	52 648	SD
Placebo	23 314	7.105	SD	SD	SD

15 years, discounted (3%)					
Treatment	**Cost (€)**	**QALYs**	**ICER (€) compared to**		
			Lowest Cost	**Next lowest cost**	**Relevant alternative**
Venlafaxine	23 423	10.644	-	-	-
Escitalopram	23 531	10.638	SD	SD	SD
Vortioxetine	23 643	10.640	SD	54 222	SD
Placebo	25 929	10.439	SD	SD	SD
25 years, discounted (3%)					
Treatment	**Cost (€)**	**QALYs**	**ICER (€) compared to**		
			Lowest Cost	**Next lowest cost**	**Relevant alternative**
Venlafaxine	24 701	16.101	-	-	-
Escitalopram	24 817	16.095	SD	SD	SD
Vortioxetine	24 929	16.097	SD	55 212	SD
Placebo	27 358	15.882	SD	SD	SD

Table A3.3. *Unequal discount rates: 4% costs, 1.5% effects*

Discount Rates: 4% Costs, 1.5% effects, 20-year time horizon					
Treatment	**Cost (€)**	**QALYs**	**ICER (€) compared to**		
			Lowest Cost	**Next lowest cost**	**Relevant alternative**
Venlafaxine	23 252	15.593	-	-	-
Escitalopram	23 359	15.586	SD	SD	SD
Vortioxetine	23 470	15.588	SD	55 725	SD
Placebo	25 738	15.359	SD	SD	SD

Table A3.4. *Equal drug acquisition costs*

Drug Acquisition Costs all equal to least costly alternative (Escitalopram)					
Treatment	**Cost (€)**	**QALYs**	**ICER (€) compared to**		
			Lowest Cost	**Next lowest cost**	**Relevant alternative**
Venlafaxine	23 454	13.518	-	-	-
Escitalopram	23 581	13.511	SD	SD	SD
Vortioxetine	23 591	13.513	SD	4 827	SD
Placebo	26 940	13.303	SD	SD	SD

Table A3.5 *Venlafaxine price increase to equal total costs with vortioxetine*

Venlafaxine cost increased from €17.63 per cycle to €182.04					
Treatment	**Cost (€)**	**QALYs**	**ICER compared to**		
			Lowest Cost	**Next lowest cost**	**Relevant alternative**
Escitalopram	26 071	13.511	-	-	-
Venlafaxine	26 138	13.518	10 467	-	10 467
Vortioxetine	26 138	13.513	32 927	0	0
Placebo	26 940	13.303	SD	SD	SD

Table A3.6 *Venlafaxine price increase to increase ICER with vortioxetine to €30 000*

Venlafaxine cost increased from €17.63 per cycle to €277.51					
Treatment	**Cost (€)**	**QALYs**	**ICER compared to**		
			Lowest Cost	**Next lowest cost**	**Relevant alternative**
Escitalopram	27 017	13.511	-	-	-
Vortioxetine	27 058	13.513	20 172	-	20 172
Venlafaxine	27 189	13.518	26 876	30 000	30 000
Placebo	26 940	13.303	SD	SD	SD

Appendix 5: Validity Tests

Adapted from Buyukkaramikli et al (2018)

Test	Expected result	Outcome achieved?
Pre-analysis calculations		
Does the probability of an event, derived from an OR/RR/HR and baseline probability, increase with higher OR/RR/HR?	Yes	Yes
Event-state calculations		
Calculate the sum of the number of patients at each health state	Sum up to cohort size	Yes
Check if all probabilities and number of patients in a state are greater than or equal to 0	Yes	Yes
Check if all probabilities are smaller than or equal to 1	Yes	Yes
Compare the number of dead (or any absorbing state) patients in a period with the number from previous periods	Should be larger	Yes
Set all utilities to 1	QALYs should equal to LYs	Yes
Set all utilities to 0	No utilities should be accumulated	Yes
Set all costs to 0	No costs should accumulate	Yes
Set the effectiveness-, utility-, and safety-related inputs for all treatment options equal	Same LYs and QALYs should be accumulated for all treatments at any time	Yes
In addition to above, set all cost-related inputs equal	Same costs, LYs and QALYs should accumulate	Yes
Calculate the number of patients entering and leaving a tunnel state throughout the time horizon	Number entering = number leaving	Yes
Results calculation		
Undiscounted results greater than the discounted results	Yes	Yes
Do the total life-years, QALYs, and costs decrease if a shorter time horizon is selected?	Yes	Yes
Check the discounted value of costs/QALYs after 2 years	Discounted value = undiscounted/$(1+r)^2$	Yes
Set discount rates to 0	Discounted results = undiscounted results	Yes
Set discount rates to a higher value	Total discounted results should decrease	Yes
Set discount rate of costs/effects to an extremely high value	Total discounted results should be similar to the discounted results in the first cycles	Yes
Uncertainty analysis		
Standard error and not standard deviation used in sampling	Yes	Yes
Lognormal/gamma distribution for HRs and costs/resource use	Yes	Yes
Beta for utilities and proportions/probabilities	Yes	Yes

Normal for other variables where samples do not violate the requirement to remain positive when appropriate	Yes	Yes
Check PSA output mean costs, QALYs and ICER compared with the deterministic results. Is there a large discrepancy?	No (in general)	No
If you take new PSA runs from the Microsoft Excel model do you get similar results?	Yes	Yes
Does the PSA cloud demonstrate an unexpected behavior or have an unusual shape?	No	No
Is the sum of all CEAC lines equal to 1 for all WTP values?	Yes	Yes
Check if sensitivity analyses include any parameters associated with structural uncertainty (e.g. discount rates, time horizon)	No	No
Is EVPI larger than all individual EVPPIs?	Yes	Yes

Appendix 6: Accordance with Observational Data

Observations taken from the introductory/background sections of this thesis.

Observation	Source	Prediction of model	Notes
50% risk of relapsing after first depressive episode	NCCMH, 2010	Approx. 70% enter 2nd-line treatment	Observation includes all MDD patients, prediction only severe cases. Prediction includes those dropping out due to AEs
Average duration of a depressive episode 30 weeks	Ferrari et al, 2013	Approx. 26% recovered and 24% in remission at week 32	Summary statistics such as averages and medians cannot be derived directly from the model.
70% of primary care patients reached full remission during a 5-year follow-up	Riihimäki et al, 2014	Approx. 66% recovered by year 5, additional 11% in remission (3rd line)	Severe cases have a poorer prognosis; however, model assumes perfect adherence, which improves prognosis.
Median weeks to remission: 20	Riihimäki et al, 2014	Approx. 53% reach remission within 24 weeks	Summary statistics such as averages and medians cannot be derived directly from the model.
Average time to remission between five and six months	Heiskanen et al, 2011	See above	See above
On average, more than 60% of patients show a response to treatment within 4-6 weeks	Heiskanen et al, 2011	Approx. 43% show a response to first-line treatment	The predicted values exclude people who may experience a response but drop out due to AEs
On average, 50% of patients are relieved from all symptoms within 4-6 weeks	Heiskanen et al, 2011	Approx. 34% in remission after first cycle	Prediction only includes severe cases, who face a poorer prognosis
84% are depression-free after 11-years	Markkula et al, 2015	Approx. 87% have recovered after 11 years	Severe cases have a poorer prognosis; however, model assumes perfect adherence, which improves prognosis.