



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e. g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Epigenetic and Population-Based Approaches to Studying the Association Between Socioeconomic Status and Depression Risk

Anders Jespersen



THE UNIVERSITY
of EDINBURGH

Doctor of Philosophy

College of Medicine and Veterinary Medicine

The University of Edinburgh

2022

Abstract

Aims

Depression is a common mental disorder and the leading cause of disability world-wide (GBD 2018). With heritability measures around 37% (Sullivan, Neale, & Kendler, 2000), most of the risk for developing depression is of a likely environmental origin. Socioeconomic status has been repeatedly associated with depression and its many comorbid diseases (Gold et al., 2020) in studies that differ widely in their age range and geographical coverage. As a result, estimates of depression risk stratified by SES are imprecise and not easily generalisable. The biological mechanisms underlying the association between SES and depression are similarly poorly understood, although differences in alcohol consumption, smoking behaviour, and obesity are widely reported. Inflammatory mechanisms have also been implicated in both SES and depression.

The aim of this thesis is to establish whether SES is associated with depression using the broadest possible evidence-base, to provide an estimate of depression risk linked to low SES, and to explore its underlying biological mechanisms.

Methods

I conducted a systematic review and meta-analysis of studies investigating the association between clinical depression and SES. There were no restrictions on country of origin, age of participants or publication date. The 140 included studies (N = 329,005, N of cases = 48,158) were meta-analysed (dichotomised into low and high SES) and reviewed, and gross national income per capita was used to interrogate differences in risk estimates and between-study heterogeneity.

To investigate the biological mechanisms associated with low SES I performed a methylome wide association study (MWAS) on the Scottish Index of Multiple Deprivation (SIMD, a continuous measure) using the single largest DNA methylation (DNAm) dataset, Generation Scotland. Estimates from the SIMD MWAS were used to predict Index of Multiple Deprivation (IMD) in the Avon Longitudinal Study of Parents and Children (ALSPAC) as a replication analysis.

I then performed an MWAS on the interaction of inflammatory disease and depression to further interrogate the findings from the SIMD MWAS and connect these findings to a possible low-SES subtype of depression. For both MWASs I performed differentially methylated point (DMP) and region (DMR) analysis, functional mapping and annotation of the differentially methylated genes, and gene set enrichment analysis.

Results

The systematic review and meta-analysis revealed an increased risk of depression associated with low SES (odds ratios from 1.57 to 1.95). The results were consistent across cultures, and SES measures. Low income was associated with the highest risk of depression. There was considerable between study effect size heterogeneity ($I^2 = 81-87\%$) in almost all meta-analysed outcomes.

The MWAS of SIMD returned 15 epigenome wide significant DMPs, and 449 DMRs. Functional annotation and gene enrichment analysis revealed bone health, cardiovascular diseases, inflammatory diseases, and major depression as associated phenotypes.

The inflammatory depression MWAS showed no epigenome wide significant DMP but 8 DMPs reached a less rigorous ‘discovery threshold’ and 7 epigenome wide significant DMRs were found. Genetic functional annotation and gene enrichment analysis implicated inflammatory markers, inflammatory diseases, response to SSRIs, liver and kidney function, and gluconeogenic pathways.

Conclusion

Findings from this thesis show that low SES and risk of clinical depression are positively associated. Longitudinal studies hint at a low SES to MDD direction of causality, but the evidence is only suggestive, and it is likely to be a bidirectional relationship. Low SES is associated with many physical measures and conditions even after accounting for smoking, alcohol consumption, and BMI. Low SES is associated with a range of inflammatory markers and chronic mental and physical diseases. The biological mechanisms underlying the link between SES and depression may be related to systemic chronic inflammation. Possible sources for this systemic inflammation include several frequently comorbid chronic health problems and

environmental toxins. These findings provide further impetus to investigate the mechanisms underlying the depression-SES association and develop interventions to mitigate the many health risks associated with low SES.

Lay Summary

Depression is a common mental disorder and the main cause of disability in the world. Some of the risk of getting depression comes from our genes but most of the risk comes from the environment we grow up in and live in. One of the things that might have an impact on the risk of depression is our socioeconomic status in society. Socioeconomic status, SES for short, is made up of how long our education was, how much money we earn and what kind of job, if any, we have. The relationship between a person's SES and their risk of getting depression is not well understood. In this PhD I wanted to measure how much more at risk of depression someone with a low SES is, compared to those with a high SES. I also wanted to find out what might happen in the body of someone with a low SES that make them more likely to get depression.

To answer these questions, I first gathered all the scientific studies that had looked at SES and depression from across the globe. By combining each of their counts of how many depressed people had a low or high SES and how many non-depressed people had low or high SES, I got a meta-result for all the people in all the studies across all the countries. That analysis showed that people with low SES are more likely to be depressed no matter where in the world you live.

Next, I looked at a part of the DNA called DNA methylation or DNAm for short. While DNA carries the information necessary to create any of the proteins that make up our bodies, the DNA methylation sits on top of the DNA making sure which bits of information are read at the correct time. Our DNAm is changed by the environments we live in and the things we do in our lives such as what we eat, if we get ill, if we exercise, or maybe start smoking. That way DNAm can work like a logbook of the things we have been exposed to in our environment and can tell us something about the way our environment impact us biologically. First, I compared DNAm from people living in low SES environments with DNAm from people living in high SES environments. I found that the DNAm from people living in low SES was different at places that had earlier been linked to the health of bones, heart diseases, inflammatory diseases, and depression.

Because the results involved low SES, depression, and inflammatory diseases I wanted to find out if these three things had anything biological in common that could

mean it was a special form of depression. So, I looked for differences in the DNAm of people who had both depression and an inflammatory disease like diabetes, asthma, or arthritis. The results were not clear, but the DNAm was different at places that had earlier been linked to toxins like phthalates and air pollution, how we metabolise food but also to the way our kidneys (helps remove toxins from our body) and liver (helps create the nutrients our body needs) function.

The results from my research show that people who have a low SES have a higher risk of getting depression. We still do not know why that is the case but it might be linked to differences in life style, things in our immediate environment and other diseases such as diabetes, asthma and arthritis.

Declarations

I declare that this thesis is composed by myself and the work is entirely my own work except where otherwise stated.

The work in this thesis has not been submitted for any other degree or professional qualification.

Anders Jespersen

May 2022

Acknowledgements

First of all, I would like to thank my PhD supervisors Professor Andrew McIntosh, Dr Heather Whalley and Professor Rebecca Reynolds. They have taught me so much and their support and guidance have been absolutely instrumental for this thesis. I would also like to thank my thesis chair Prof Sue Fletcher-Watson who has been a great source of support and inspiration.

I want to say a big thank you to fellow PhD students and postdocs in the research group who have helped me so much both academically and socially. In particular I want to thank Niamh, Alex, Melissa, Toni, Melisa, and Mark for all their help.

A big thank you goes out to the Wellcome Trust for funding this PhD but also for constantly striving to make UK research better, both for the research and for the researchers. The Wellcome Trust Translational Neuroscience PhD programme at the University of Edinburgh deserves a huge thank you. Thank you to everyone involved for creating such a supportive and vibrant academic environment. In particular I want to thank Jane Haley, David Price, and of course my Translational PhD pals Bex, Caoimhe, Fenia, Rana, and Rob. I also want to thank Stephen Lawrie without whom I would not have heard about the programme.

I would like to thank my family for supporting me throughout my education and making it possible for me to pursue this PhD. I also want to thank my dog Alba for always being a positive presence and always being up for a walk when I needed to clear my head.

Finally, I want to say an enormous thank you to Dr Alexandra Mitchell. Without you I would have never made it through this PhD. Your never-ending support and academic input have been absolutely invaluable.

Table of Contents

<i>Abstract</i>	<i>i</i>
Aims	i
Methods	i
Results	ii
Conclusion	ii
<i>Lay Summary</i>	<i>iv</i>
<i>Declarations</i>	<i>vi</i>
<i>Acknowledgements</i>	<i>vii</i>
1 Introduction	1
1.1 Depression	1
1.1.1 Prevalence.....	1
1.1.2 Diagnosis	2
1.1.3 Symptom heterogeneity	3
1.1.4 Age of onset.....	4
1.1.5 Comorbidity	4
1.1.6 Treatment.....	5
1.1.7 Societal & Economic cost.....	7
1.1.8 Genetic risk factors	8
1.1.9 DNA methylation	10
1.1.10 Environmental risks	11
1.2 Socioeconomic status	15
1.2.1 SES and physical health	16
1.2.2 SES and mental health	18
1.3 Inflammation	20
1.4 Summary	23
1.5 Aims	24
2 Socioeconomic Status and Depression – A Systematic Review	26
2.1 Introduction	26

2.2	Methods	30
2.2.1	Inclusion and exclusion criteria	30
2.2.2	Search strategy	32
2.2.3	Data collection	32
2.2.4	Analysis	34
2.3	Results	36
2.3.1	Study inclusion	36
2.3.2	Missing data and contact with authors	38
2.3.3	Study population	39
2.3.4	Dichotomised meta-analyses	40
2.3.5	Ordinal three group meta-analyses	48
2.3.6	Continuous data meta-analysis	54
2.3.7	Longitudinal and narrative review	55
2.4	Discussion	64
2.4.1	Main findings	64
2.4.2	Strengths and limitations	68
2.4.3	Future work	70
3	<i>DNA Methylation Signatures of Area-Based Deprivation</i>	72
3.1	Introduction	72
3.2	Methods	74
3.2.1	Phenotypes	74
3.2.2	DNA methylation	76
3.2.3	Statistical methods	77
3.2.4	Differently Methylated Region Analysis	78
3.2.5	Replication analysis	78
3.2.6	Training dataset	79
3.2.7	Testing Datasets	79
3.3	Results	80
3.3.1	Demographics	80
3.3.2	Differentially Methylated Positions	81
3.3.3	Differentially Methylated Regions	86
3.3.4	Lasso prediction of SIMD/IMD quintile	88
3.4	Discussion	89
3.4.1	Strengths and limitations	91
3.4.2	Future directions	93
4	<i>DNA Methylation Signatures of Inflammatory Disease and Depression</i>	95

4.1	Introduction	95
4.2	Methods	97
4.2.1	Phenotypes	97
4.2.2	DNA methylation	98
4.2.3	Statistical methods	99
4.3	Results	100
4.3.1	Inflammatory Disease and Depression	100
4.3.2	Differentially Methylated Positions	101
4.3.3	Differentially Methylated Regions	105
4.3.4	Replication	107
4.4	Discussion	107
5	<i>Discussion.....</i>	<i>115</i>
5.1	Introduction	115
5.2	Summary of main findings.....	117
5.2.1	Low socioeconomic status, in particular low income, is associated with increased risk of depression across cultures.	117
5.2.2	Area-based deprivation is associated with widespread differences in DNAm patterns, most of which are associated with inflammatory processes.	119
5.2.3	Exploring biological mechanisms connecting inflammatory diseases and depression ..	124
5.3	Strengths and limitations	127
5.4	Future research	130
5.5	Conclusions.....	131
	<i>Bibliography</i>	<i>133</i>
	Systematic Review References.....	176
	<i>Appendix 1: FUMA analyses of SIMD EWAS DMP genes.....</i>	<i>194</i>
	<i>Appendix 2: FUMA analyses of SIMD EWAS DMR genes.....</i>	<i>200</i>
	<i>Appendix 3: SIMD LASSO coefficients</i>	<i>212</i>

1 Introduction

This thesis covers investigations into the association between depression and socioeconomic status, as well as the biological pathways underlying this association. The following sections of this introduction will cover background information of symptomatology and epidemiology of depression, socioeconomic status as a risk factor for poor health, DNA methylation as a possible tool for investigating biological impacts of environment exposures, and the role of inflammation in mental and physical health.

1.1 Depression

1.1.1 Prevalence

Depression is the most common mental health disorder, and it is the leading cause of disability worldwide (W H O, 2021). It is estimated that 322 million people suffer from depression, resulting in a global prevalence of around 4.4% (W H O, 2021). It is argued that the true burden of depression (and indeed other mental health disorders) is difficult to estimate as it is often comorbid with other mental health disorders, chronic pain, self-harm and suicide making it difficult to disentangle the source of the disability and burden (Vigo, Thornicroft, & Atun, 2016). There is a clear sex difference in the prevalence of depression, women are more likely to suffer from depression with a global prevalence of 5.1% while men have a prevalence of 3.6%. The difference in risk of depression between sexes seems to be universal across cultures (Hopcroft & Bradley, 2007).

Prevalence rates of depression vary across countries and world regions, ranging from 2.6% in men in the Western Pacific region to 5.9% in women in the African region (W H O, 2021). It is unclear whether differences in prevalence rates are significantly different by country GDP. A review of epidemiological data found a higher life-time prevalence in low and middle (compared to high) income countries but no difference in 12-month prevalence (Kessler & Bromet, 2013). However, environmental risk factors were found to be relatively consistent across cultures.

1.1.2 Diagnosis

There are several different ways of diagnosing depression both culturally and throughout history. Diagnostic techniques of depression are varied and often differ across diagnostic setting (clinical, community, population based) and country, in particular because of differences in languages. The Diagnostic and Statistical Manual of Mental Disorders-V (DSM-V) states that to qualify for a diagnosis of major depressive disorder (MDD) you must present with at least 5 symptoms of depression for 2 weeks or more, at least one of which must be either depressed mood or reduced interest or pleasure (anhedonia). The remaining symptoms from which the threshold of 5 must be met include weight loss or gain, insomnia or hypersomnia, psychomotor agitation or retardation, fatigue, feelings of worthless or excessive guilt, decreased concentration, and thoughts of death/suicide. The symptoms must cause significant impairment to everyday functioning, and they should not be attributable to any physiological effects or substances or other medical conditions, including schizophrenia, bipolar disorder and related psychiatric disorders.

One of the most common ways to diagnose depression in a clinical setting is by using the Structured Clinical Interview for DSM-5 (SCID), a structured interview designed to systematically and reliably measure the presence or absence of core depressive symptoms (First, 2016). However, there are many other ways of assessing the presence of clinical depression. More than 280 measures of depression or depressive symptom severity have been developed since 1918 (Santor, Gregus, & Welch, 2006). One of the most widely used measures (Sayer et al., 1993) is the Beck Depression Inventory (BDI) (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). It is one of the more common screening tools and it is also used to measure the severity of depressive symptoms. Another widely used screening tool is the Centre for Epidemiologic Studies Depression Scale (CES-D). It is designed to assess depression in the general population and has proven a consistent reliability and validity across some cultures and age groups.

1.1.3 Symptom heterogeneity

Depression is unequivocally clinically heterogeneous. With nine core symptoms of depression, three of which can go in opposite directions (weight loss/gain, insomnia/hypersomnia, psychomotor agitation/retardation) there are 227 possible combinations of symptoms leading to an MDD diagnosis (Zimmerman, Ellison, Young, Chelminski, & Dalrymple, 2015). Symptoms of MDD tend to cluster in some combinations more than others, suggesting there may be subtypes of depression with similar presentations and potential aetiologies (Zimmerman et al., 2015).

It is argued that viewing MDD as one disorder has been beneficial in terms of reducing stigma, increasing awareness and public investment, but may prove a hindrance to the advancement of diagnosis and treatments (Goldberg, 2011). Instead, Goldberg (2011) argues that we need a stratification of depression into subtypes to succeed in increasing treatment efficacy. Similarly (Fried & Nesse, 2015) suggest that specific depression symptoms are distinct from one another and differ in terms of the biological mechanisms and environmental risk factors involved. And as a result, we may overlook important biomarkers if we keep using aggregate sum scores of symptoms when researching depression.

A recent systematic review by (Margaroli, Calderon, & Bonanno, 2021) found that fatigue and depressed mood tend to form two separate symptom clusters, and much research over the past couple of decades have focused on teasing apart depression into subtypes. These studies are often based on clustering of symptoms, comorbid diseases, genetic markers, or environmental stressors. A recent review of depression biomarker literature (Strawbridge, Young, & Cleare, 2017) found support for five distinct biomarker systems of depression that could point to possible depression subtypes. The five groupings are; inflammatory, neuroendocrine, growth factor, neurotransmitter and metabolic. Neuro endocrine processes such as the HPA axis and hypercortisolism, and neurotransmitters such as 5-HT, GABA and histamine have been researched for many decades. More recently, and increasingly, inflammatory markers and diseases, growth factors (BDNF, GDNF, IGF-1) and metabolic processes (insulin, glucose, lipids, albumin) have been implicated in

depression. Nevertheless, in spite of many decades of research, MDD remains heterogeneous and is yet to be consistently dissected into clinical subtypes or stratified into subgroups with distinct aetiologies.

1.1.4 Age of onset

The prevalence and prognosis of depression also differs across the lifespan. With an average age of onset of 25.5 years, depression has a later age of onset compared to many other mental health disorders (Kessler et al. 2013) such as schizophrenia, bipolar disorder and neurodevelopmental conditions. The prevalence of depression increases over the life course and peaks in older adulthood (World Health Organization, 2017). This may be related to the fact that an early onset of depression (< 25 years) is associated with a greater risk of recurrent depression as opposed to a later onset of depression, thereby accumulating the incidence of depression at any given age. Similarly, there is evidence that an adolescent onset of depression is associated with increased risk of a particularly severe depression symptomatology including worsening cognitive functioning and risk of self-harm and suicide (Wilson, Hicks, Foster, McGue, & Iacono, 2015; Zisook et al., 2007).

Early work on depression in old age identified a possible link between cerebrovascular disease and a later onset of depression (Alexopoulos et al., 1997). This was later referred to as vascular depression, signified by cerebrovascular damage and an age of onset later than 65 years old (Aizenstein et al., 2016). Late onset of depression has also been linked to increased risk of developing dementia and has even been proposed as a prodromal phase or early marker for dementia (Bennett & Thomas, 2014).

1.1.5 Comorbidity

Major depression is associated with many comorbid diseases and disorders. The prevalence of depression in inpatients (12%) and general outpatients (27%) is many-fold higher (Walker et al., 2018) than the prevalence of depression in the general population (4.4%).

Depression is highly comorbid with most cancers (Mitchell et al., 2011). The high comorbidity may in part be due to pain and fatigue severity as well as increased levels of stress (Gold et al., 2020). Across the board depression shows increased prevalence in people with cardiovascular disease, metabolic diseases such as diabetes type 1 and type 2, stroke, multiple sclerosis, and inflammatory diseases in general (Gold et al., 2020).

There are many possible reasons for the high comorbidity rate of depression with a wide range of physical health problems. Some of the causes could be shared genetic aetiology, shared environmental risks, a change in health behaviours or as a side-effect of medication such as corticoid steroids. Systemic inflammation as a cause or a consequence of a physical disease has also been suggested as the reason for the high comorbidity (Enache, Pariente, & Mondelli, 2019).

Depression is also highly comorbid with other mental disorders. A large-scale register based cross sectional analysis of 6.3 million individuals diagnosed with depression and 25 million age matched controls found that almost all mental disorders were at least twice as prevalent in people with depression and at least 64% of depression cases had a comorbid mental disorder (Steffen, Nübel, Jacobi, Bätzing, & Holstiege, 2020). Anxiety disorders were most commonly comorbid, followed by substance use disorders and schizophrenia. Receiving a diagnosis of a mood disorder (depression) before age 20 years was associated with increased risk (30.6% for men and 38.4% for women) of developing neurotic disorders within 5 years (Plana-Ripoll et al., 2019). This level of comorbidity introduces further complexity to the study of depression impacts and aetiologies.

1.1.6 Treatment

Treatment of depression can be broken into two broad categories: physical treatments and psychological therapies. Physical treatments include drugs, electroconvulsive treatment, psychosurgery, transcranial magnetic stimulation, and vagal nerve stimulation. The most common treatment for depression is via anti-depressant medication (Jobst et al., 2016) around 70% of adults with a depression diagnosis in the US receive anti-depressant medication. Monoamine oxidase inhibitors (MAOIs) and tricyclic antidepressants (TCAs) were some of the earliest anti-

depressant medications available, both were discovered largely by chance (Ramachandrai, Subramanyam, Bar, Baker, & Yeragani, 2011). Due to strong side-effects more selective anti-depressant medications were developed such as Selective Serotonin Reuptake Inhibitors (SSRIs) and Serotonin/Noradrenaline-Reuptake Inhibitors (SNRIs). SSRIs presently are the most common type of antidepressants; they work by reducing the reuptake of serotonin thereby increasing the extracellular levels of serotonin which in turn increases monoamine neurotransmission. The exact mechanisms behind the anti-depressant effects of increasing available cerebral serotonin levels are not well understood. However, a prevailing theory suggests increased levels of serotonin can cause downstream effects leading to an increase in neuroplasticity and ultimately a change in symptoms (Harmer, Duman, & Cowen, 2017).

Antidepressants are most often prescribed as monotherapy, one antidepressant at a time. However, as symptom severity increases so does the proportion of patients on a combination therapy of several antidepressants: 8% for mild, 16% for moderate, and 27% for severe depressive symptoms (Cascade, Kalali, & Blier, 2007).

A recent meta-analysis has shown that antidepressant medication is highly effective for most patients (Cipriani et al., 2018). However, not everyone responds well to anti-depressant medication. Treatment resistant depression refers to the failure to produce a clinically meaningful effect after two consecutive treatments with medication from different pharmacological classes at an adequate dose. A clinically meaningful effect most often refers to a 50% reduction in symptoms. It is estimated that between 18% and 30% of people with a diagnosis of depression are suffering from treatment-resistant depression (Al-Harbi, 2012; Jaffe, Rive, & Dene, 2019). Side-effects from antidepressant medication are very common, it is estimated that 55% of patients on anti-depressants experience some sort of unwanted side effect (Hu et al., 2004). Common side-effects include fatigue, sexual dysfunctions, gastrointestinal problems, weight gain, insomnia, tremors, and apathy (Kelly, Posternak, & Alpert, 2008).

The most common non-pharmacological treatment for depression is cognitive behavioural therapy (CBT). In a meta-analysis of 140 different treatments for depression CBT was the only non-pharmacological treatment with a reduction in depressive symptoms similar to that of anti-depressant medication

(Gartlehner et al., 2017). CBT was developed in the 1970's (Beck, 1970) and employs systematic change of behaviour, belief change, and cognitive restructuring (Sudak, 2012). CBT have been found to have a comparable effect to anti-depressant medication in terms of symptom reduction (DeRubeis, Gelfand, Tang, & Simons, 1999) and some studies suggest that patients receiving CBT are less likely to relapse after treatment has ended. In a randomised control trial (DeRubeis et al., 2005) 50% of patients who had received CBT had not relapsed 12 months after treatment ended. Less than 10% of the patients receiving anti-depressant medication had avoided a relapse 12 months after treatment ended. A recent meta-analysis found a small but significant improvement in depressive symptoms when patients received a combination of psychopharmacological therapy and psychotherapy compared to any one standalone treatment (Cuijpers van Straten, Warmerdam and Andersson 2015; Dunlop et al., 2019). In most countries CBT is not often offered to acute patients as part of standard healthcare programme as it is a time consuming and costly treatment.

1.1.7 Societal & Economic cost

With a high prevalence, high rates of relapse, and disabling symptoms, it is no surprise that depression has a significant societal and economic impact. People with a mood disorder have been found to have a six-to-eight-year decrease in life expectancy (Plana-Ripoll 2019). Similarly, a meta-analysis found an association ($OR = 1.52$, $CI = 1.45-1.59$) between excess mortality and depression (Cuijpers et al., 2014). Part of this lower life expectancy and excess mortality is due to an increased risk of suicide. Up to half of all suicides world-wide are associated with depressive or mood disorders (Bachmann, 2018), resulting in a considerable cost to the individual, their families, and society as a whole.

Early onset of depression is associated with high personal and societal costs such as lower educational attainment, negative parenting behaviours, poorer marital relationship quality, and increased risk of being the victim of marital violence (Kessler, 2012).

The economic burden of depression, to society is vast. It is estimated that between 2005 and 2010 depression was associated with a cost of \$210.5 billion in the

United States (Greenberg, Fournier, Sisitsky, Pike, & Kessler, 2015). The cost of depression in Europe has been estimated at €118 billion in 2004 (Sobocki, Jönsson, Angst, & Rehnberg, 2006) although it is argued that the real cost is much higher (Ekman, Granström, Omérov, Jacob, & Landén, 2013). A few studies have investigated whether early and increased interventions for depressed people at out-patient clinics had an economic benefit (Rost, Smith, & Dickinson, 2004; Wang et al., 2007). The interventions, which included enhanced care and structured telephone cognitive behavioural psychotherapy, saw increased productivity and decreased absenteeism. Overall, the cost of the early intervention and screening programmes were far outweighed by the economic benefits of reducing the burden of depression.

1.1.8 Genetic risk factors

Observations of higher rates of depression in some families than the general population and in the twin of an affected individual, in general, and in monozygotic twin in particular, cemented the notion that depression is probably heritable. Further studies showed that the risk of depression in the co-twin of an affected twin was higher when the siblings were identical, provided further evidence for a partly genetic basis of depression. Twin and family studies may be used to estimate the proportion of variation in liability within a population due to genetic differences. This value is frequently called the heritability of a trait. A meta-analysis of twin-studies estimated the heritability of depression to be between 34% and 60% (Polderman et al., 2015).

Given the general acceptance that depression is at least partly caused by genetic factors, researchers initially started conducting candidate gene studies in which a single or a few individual genes were investigated based on a priori assumptions (McInnes & Freimer, 1995). One of the more famous examples of this is the serotonin transporter gene 5-HTT which was hypothesised to be involved in depression risk due to the antidepressant effects of SSRIs (Heils et al., 1996). However, many of the studies failed to replicate and a recent, and adequately powered, analysis of several previously reported depression candidate genes failed to find any significant association between their polymorphisms and depression (Border et al., 2019). Early exploratory studies investigating the entire genome; genome wide association studies

(GWAS), failed to find any genetic variants associated with depression when applying stringent corrections for repeated testing (Sullivan et al., 2009; N R Wray et al., 2012). Not until a female only case-control matched sample of 10,000 individuals or a cohort exceeding 300,000 participants were used, were genome wide significant single nucleotide polymorphisms (SNPs) identified (Cai et al. 2015, Hyde et al. 2016), an indication of just how small the effect of each SNP is. By then it was increasingly clear that that depression is a highly polygenetic disorder with thousands of genetic variants each exerting a very small effect on the risk of developing depression.

The most up to date genome wide association study on common genetic variants associated with depression found 102 independent variants (Howard et al., 2019). Pathway analysis revealed an enrichment of genes involved in synaptic structure, neurotransmission, and cell types specific to frontal cortex and anterior cingulate cortex.

Interestingly, none of the candidate genetic variants associated with the serotonergic system reached genome wide significance in GWASs of depression to date (Dunn et al., 2015; Howard et al., 2019). It is suggested that the biological pathways of depression disorder and depression treatment are in fact separate but intertwined through intermediary biological processes.

There is considerable overlap in genetic variants associated with depression and other disorders and diseases. (Howard et al., 2019) found genetic correlations between depression and schizophrenia, bipolar disorder, coronary artery disease, Crohn's disease, and body fat, but also age at menopause and even college completion showed a genetic correlation with depression.

Early reviews (Sullivan 2000) of the genetic risk of depression acknowledged that while there is a definite genetic contribution to the risk of depression there is no doubt that the interplay between an individual's environment and their genetic make-up, and even the genetic makeup of those in their immediate environment (non-transmitted genetic effects), all contribute to the overall risk of depression. Maybe even more so than for other major psychiatric disorder – therefore it is important that we have an understanding of the impact of environment on the individual. This is often problematic, and not easy to quantify. However DNAm might provide a means of exploring the biological impact of the lived environment.

1.1.9 DNA methylation

At the intersection of genetic and environmental risk factors sits DNA methylation (DNAm). DNAm is a type of epigenetic change as it “sits on top” of the genome and regulates which genes are transcribed when (Tost, 2009). DNAm refers to the addition of a methyl group to the DNA strand, usually on the phosphate bond between a cytosine and guanine amino acid, hence described as a CpG site. The addition of methylation will usually silence the transcription of a gene if it is located at the promoter region of that gene (Newell-Price, Clark, & King 2000).

DNAm levels changes with environmental exposures (Fraga et al., 2005) and these changes are passed on to daughter cells. Direct causal evidence of environmental exposures impacting DNAm and consequently changing gene expression in humans is scarce and difficult to isolate (Barker, Walton, & Cecil, 2018). Similarly, it has proven difficult to investigate whether, theoretically possible, transgenerational epigenetic inheritance can be linked to differences in health risk (Heard & Martienssen, 2014). However, differences in DNAm have previously been linked to smoking, alcohol consumption, obesity, cancers, and depression (Miruna C Barbu et al., 2019; Chen et al., 2021; Joeannes et al., 2016; Lohoff et al., 2021; Spainhour, Lim, Yi, & Qiu, 2019). Crucially these changes are mechanistically specific, and, in some cases, there is evidence of a dose-dependent relationship between environmental exposure and change in DNAm (McCartney et al., 2018). This presents an intriguing opportunity to use DNAm levels as a biological footprint of environmental risk or indeed a ‘logbook’ of biological mechanisms impacted by the environment.

DNAm levels have also been used as disease biomarkers. Recently DNAm scores have been calculated for over 100 blood protein levels and these scores were used to predict disease risk (Gadd et al., 2022). Deviations between chronological and biological age has also been calculated using so called epigenetic clocks, based on differences in DNAm (Hannum et al., 2013; Horvath, 2013). DNAm derived plasma protein scores have even been used to predict lifespan (Lu et al., 2019). DNAm has also been employed to calculate longitudinal scores for otherwise volatile inflammatory markers such as CRP and IL-6 and have in turn been associated with differences in cognitive ability and cognitive functioning (Stevenson et al., 2018, 2020). More recently, work from our group published in the

last year of this thesis has shown differences in DNAm associated with antidepressant-medication and genetic risk of depression, both of which implicated immune processes (Miruna C Barbu et al., 2021; Shen et al., 2022).

1.1.10 Environmental risks

With an estimated heritability of 37% a considerable contribution to the risk of developing depression is of a likely environmental origin.

One of the earliest environmental risks of depression is childhood trauma. A large-scale meta-analysis found that childhood maltreatment was associated with increased risk of depression and worsening depressive symptoms (Humphreys et al., 2020). All types of childhood maltreatment were associated with increased risk of depression, but emotional abuse and emotional neglect had the strongest association. One review suggests that half of global depression cases could be at least partially attributable to childhood maltreatment (M. Li, D'Arcy, & Meng, 2016). In fact, one of the earliest categorisations of the risks of depression, from 1621 (Brink 1979), suggested that a traumatic childhood, the loss of a loved one, poverty and loneliness could cause depression.

In general, stressful life events, such as moving home or country (Bhugra & Ayonrinde, 2004) have been associated with depression. Most research has focused on episodic stressors (individual events) such as war, pandemic, bereavement, or amputation but also other stressful life events such as divorce, job loss, financial difficulties or moving country (Constance Hammen, 2005). It is however, important to remember that the level of stressfulness from a particular event may be different from person to person. For decades it has been hypothesised that stressful life events increase the risk of depression through increased activity of the hypothalamic–pituitary–adrenal (HPA) axis (Carroll et al., 1981).

During a stressful event activation of the central nervous system (CNS), causes the release of hormones (corticotropin releasing hormone, adrenal corticotropic hormone, and cortisol) from the hypothalamus, the anterior pituitary, and the adrenal cortex. Most impactful of those, the glucocorticoid cortisol. In turn, during the recovery phase after a stressful event the increased level of cortisol causes a negative feedback system in the hippocampus to inhibit the HPA system

(L. Jacobson & Sapolsky, 1991). Studies have consistently linked differences in HPA activity to depression (Jurueña, 2014) and levels of cortisol to severity of symptoms (C Hammen, Davila, Brown, Ellicott, & Gitlin, 1992). However, more recent reviews suggest that methodological issues may account for the severity of previous findings and while cortisol levels may be higher in depressed individuals the effect is stronger in older participants (Stetler & Miller, 2011). Furthermore, there seems to be stronger evidence for an aberrant recovery period for depressed individuals after a stressful event, leading to a longer recovery period for their HPA system (Burke, Davis, Otte, & Mohr, 2005).

Many so-called lifestyle factors have been identified as risks of depression, one of those is body mass index (BMI). A high or low BMI on its own does not necessarily confer a health risk and has also been criticised as a poor indicator of an individual's health as it does not give a good indication of an individual's body fat percentage or the distribution of said fat on the body, something that has become increasingly linked to health issues (Nuttall, 2015). However, BMI does correspond well to bodyfat percentage within sex and age groups at the group level (Flegal et al., 2009). As a result, and because of the ease of measuring it, it is often used as a proxy for obesity at a population level.

Historically findings on the association between BMI have been inconsistent. High BMI indicating extreme obesity has been associated with an increased risk of depression (OR = 1.73, CI = 1.04-2.87) after controlling for age, sex, chronic diseases, and socioeconomic status (C. Dong, Sanchez, & Price, 2004; Robert E Roberts, Strawbridge, Deleger, & Kaplan, 2002; R E Roberts, Kaplan, Shema, & Strawbridge, 2000). However, other studies have found no effects (Friedman & Brownell, 1995) or reverse effects of BMI on depression risk (Palinkas, Wingard, & Barrett-Connor, 1996). A recent Mendelian Randomisation (MR) study used genetic variants associated with BMI and depression to investigate the association and direction of causality between BMI and depression. Results revealed a strong association between increased BMI and increased risk of depression particularly in women. The results from the MR analysis were less clear but indicated that high BMI is has a causal role in increased risk of developing depression (Tyrrell et al., 2019). There is a substantial genetic overlap between BMI and Depression (Milaneschi, Simmons, van Rossum, & Penninx, 2019)

and genetic correlations haven been found between BMI and mental disorders in general and depression in particular (Bahrami et al., 2020).

Alcohol misuse is associated with increased risk of depression. A large-scale meta-analysis of the association between alcohol use disorder (AUD) and depression found a strong association between the two (Boden & Fergusson, 2011). The association was found to go either way, however, the authors highlight that epidemiological data suggest AUD is causally associated with depression. A more recent study found that alcohol consumption promotes depression, the findings were consistent across drinking behaviour; amount consumed, consumption intensity and alcohol dependence (Awaworyi Churchill & Farrell, 2017). It has been hypothesised that alcohol causes depression through pharmacological effects on the brain. Excessive alcohol consumption is thought to increase the risk of depression through the release of peptides that over-activate the HPA axis (Bruijnzeel & Gold, 2005). Conversely it is hypothesised that depression could cause increased alcohol consumption through self-medicating to alleviate symptoms (Crum et al., 2013). As a resolve to the question of direction of causality it has been suggested that a partly shared biological risk of the two disorders could explain the bi-directional relationship. This was illustrated by a recent study on Polygenic risk scores (PRS) of depression and alcohol dependence, showing a shared genetic susceptibility for the two disorders (Andersen et al., 2017).

Apart from being a major public health concern on its own, tobacco smoking also has a bidirectional relationship with depression, much like alcohol consumption. Research show that cigarette consumption is much higher in people with mental illness and while the number of smokers in the general population has decreased, the number of smokers among people with a mental illness has remained at a relatively high level (McManus, Meltzer, & Campion 2009). A meta-analysis of 148 studies found roughly half of the studies showed an association between smoking behaviour and subsequent onset of depression or anxiety, and the other half reporting initiation of smoking followed by a diagnosis of depression or anxiety (Fluharty, Taylor, Grabski, & Munafò, 2017).

Although much less well studied than other risk factors, environmental toxins have also been associated with depression (van den Bosch & Meyer-Lindenberg, 2019). Air pollution, such as carbon monoxide, particulate matter has consistently been associated with a respiratory and cardiovascular health problems

(Brook et al., 2010; Carlsen, Forsberg, Meister, Gíslason, & Oudin, 2013) as well as mental health problems (Tallon, Manjourides, Pun, Salhi, & Suh, 2017). Small particulate matter (<2.5 microns) seems especially dangerous as studies show the fine particles can reach blood circulation (most often through inhalation) and ultimately reach the brain (Calderón-Garcidueñas et al., 2015). A recent review found an association between high levels of air pollution and increased risk of suicide (Ragguett et al., 2017) Similarly in a review, 24 out of 26 studies showed a significant association between heightened exposure to air pollution and increased risk of depression (van den Bosch et al. 2019).

Studies have used animal models of mood disorders to understand the causal pathways between air pollution and depression. Results have shown a link between exposure to air pollution and increased hippocampal inflammation ($t(11) = 2.22$, $p < 0.05$) coupled with worsening depressive symptoms (Fonken et al., 2011). Similarly, a study found no increases in BDNF (brain-derived neurotrophic factor) after cycling along busy roads (Bos et al., 2011). BDNF increases brain plasticity by promoting synapse formation and increasing synaptic spine density (Yang et al., 2020), levels are normally increased after exercise and increases in BDNF have been associated with successful anti-depressant treatment whereas decreases in BDNF levels in hippocampus and prefrontal cortex have been associated with depression.

Phthalates, used in the manufacture of plastics, have also been linked to increased risk of depression (Bao et al., 2022; M. H. Jacobson et al., 2021; K.-S. Lee et al., 2018). Preclinical models have shown increased exposure to phthalates can cause dysregulation of BDNF in hippocampus (Rowdhwal & Chen, 2018) and have been found to impact the HPA axis (Hliseníková, Petrovičová, Kolena, Šidlovská, & Sirotkin, 2021). Urinary levels of phthalates (mono-n-butyl) have been associated with increased risk of depression (OR = 1.33, CI = 1.14-1.54) in adults (Shiue, 2015) and with worsening depressive symptoms in an elderly population (K.-S. Lee et al., 2018).

Heavy metal exposures, most of which happens through diet, traffic, and industry, have also been associated with depression. Some studies have found a link between risk of depression and levels of cadmium, lead, and mercury (Berk et al., 2013; Bouchard et al., 2009; Kern et al., 2014). The strongest link seems to be between mercury and depression risk. However, the findings for other heavy metals are far from clear-cut and some longitudinal studies show no relationship

between heavy metal concentrations and depression risk (C. Han, Lim, & Hong, 2016).

1.2 Socioeconomic status

Socioeconomic status (SES) is a major risk factor for depression although the biological mechanisms responsible for this association remain unknown. SES as a risk factor for depression is less well studied than genetic and biological risk factors mentioned above. SES refers to an individual's social standing or class. SES is most often measured by combining a person's educational attainment, income level and occupational category or job status (employed vs. unemployed). SES is often used as a proxy for social and material deprivation as it is easier to measure at a population level and can often be obtained from public registers. Deprivation, whether social or material, refers to the lack of benefits that in the given society, are a basic necessity. In that way one could argue that deprivation could be the result of a person's socioeconomic status.

Measures of material and social deprivation differ from country to country. A very commonly used measure in the UK is the Townsend Index of Deprivation. The Townsend Index of Deprivation is a neighbourhood measure incorporating the variables; unemployment, non-car ownership, non-homeownership, and household overcrowding (Jarman, Townsend, & Carstairs, 1991). Two other commonly used measures are the English index of multiple deprivation (IMD) and the Scottish index of multiple deprivation (SIMD) used in England and Scotland respectively. Both the IMD and the SIMD divide the country into small areas with an average of 1500 or 750 people in each area and then rank all the areas from least to most deprived based on several social factors. Those factors are income, employment, education, health, crime, access/barriers to services, and housing.

In this thesis I will be dealing with both multiple deprivation and SES. The measurement of SES can be divided into two broad categories, individual and neighbourhood measures. There are even suggestions that individual and neighbourhood SES can interact so that people with a low individual SES living in a high SES neighbourhood may be protected from some of the cardiovascular health

risks otherwise associated with low SES (Boylan & Robert, 2017). As such it can be argued that neighbourhood and individual SES are associated with mostly overlapping but also unique contributions to individual health.

1.2.1 SES and physical health

Low SES has been associated differences in risk of ill-health for a very long time (Collier, 1937). However, it is only really since the mid 1980's that SES has systematically been studied as a variable of interest and not as something to be controlled in an analysis (Adler & Ostrove, 1999). The prevailing theory was the threshold model, arguing that increasing wealth below the poverty line would result in health improvements but above the poverty line increases in wealth would make no significant changes in health. One of the main reasons for the change of attitude was the Whitehall Study (Marmot, Shipley, & Rose, 1984). The Whitehall study followed 17,000 civil servants for 10 years and showed a steep inverse relation between employment grade and mortality, across all employment grades (Marmot et al. 1984). The 10 year all cause mortality percentage was 4.7% (SE = 0.7) for administrative staff, 8.0% (SE = 0.3) for professional and executive staff, and 11.7% (SE = 0.6) for clerical staff. A more recent report from Marmot et al Marmot (2020) on health equity has shown that the differences in life expectancy between the least deprived and the most deprived areas are in fact increasing and in the 10% most deprived areas women's life expectancy has decreased from 2010 to 2018. In other words, health inequality in the UK is worsening. It is not only mortality that is affected, but healthy life expectancy (the number of years spent with good health) also shows an even steeper inverse relationship with deprivation (Marmot et al 2020).

It is argued that social drift is more likely for health problems with an early onset such as schizophrenia, however there is strong evidence for the social causation theory with childhood socioeconomic factors impacting health far into old age (Dahl & Birkelund, 1997) even after controlling for adult SES levels (Hertzman, 1999). (Warren, 2009) used longitudinal data to specifically test the hypothesis whether changes in health affected SES or whether SES affected health throughout life. They found no support for the former, SES affected health at multiple

timepoints throughout life but the reverse, health affecting SES, was not found. More specifically low SES has previously been associated with cardiovascular disease, certain cancers, diabetes, arthritis, chronic respiratory diseases (COPD, asthma, cystic fibrosis, and pulmonary hypertension), gastrointestinal diseases in children, and obesity, although the association between low SES and obesity is weakening (Adams et al., 2018; Bengtsson et al., 2005; Kyrou et al., 2020; McLaren, 2007; de Mestral & Stringhini, 2017; Sahni, Talwar, Khanijo, & Talwar, 2017; Wong, Lao, Ho, Goggins, & Tse, 2017).

The mechanisms behind these relationships are poorly understood. An early and popular theory revolves around health behaviours, arguing that health impacts associated with low SES are, in part, driven by an unhealthy lifestyle. It has been argued that the abundance of unhealthy foods, alcohol and tobacco in low SES neighbourhoods has accumulating, long lasting health impacts (Y.-C. Chuang, Cubbin, Ahn, & Winkleby, 2005; Hiscock, Bauld, Amos, Fidler, & Munafò, 2012; Reidpath, Burns, Garrard, Mahoney, & Townsend, 2002). However, the association does not seem to be universal (Macintyre, McKay, Cummins, & Burns, 2005) and in some cases low SES was associated with an increased availability of shops but a lower number of heavy drinkers compare to high SES (Pollack, Cubbin, Ahn, & Winkleby, 2005). There is strong evidence to suggest that low SES is associated with increased alcohol consumption, and a higher prevalence of smoking and obesity (Beard et al., 2019; Hiscock et al., 2012; Vazquez & Cubbin, 2020). However, mounting evidence suggest that these health behaviours far from explains all the association between SES and health. Changes in smoking behaviour, alcohol consumption and obesity rates could not explain the steepening of the inverse relationship between SES and mortality over the last few decades (Cutler, Lleras-Muney, & Vogl, 2008).

Some mechanisms are more direct, for instance lead ingestion due to poor housing can have physical and mental health consequences, amongst others depression, as described earlier (Afeiche et al., 2012; Lidsky & Schneider, 2003). Levels of air pollution are often higher in low SES neighbourhoods and has been associated with increased risk of cardiovascular disease, depression and Alzheimer's disease (van den Bosch & Meyer-Lindenberg, 2019; Brook et al., 2010; Fu & Yung, 2020). Other mechanisms can work

through differences in policy for instance lower SES is associated with poorer sick-leave benefits forcing employees to go to work ill, thereby worsening their health and increasing the risk of transmission of communicable diseases in the community (Cook, 2009). Similarly, (Nandi, Glymour, & Subramanian, 2014) found that only around half of the increased mortality risk associated with low SES could be explained by smoking, alcohol consumption and physical inactivity.

Chronic stress in early life, and its biological consequences, has been proposed as the general underlying mechanism of the SES and health association. Childhood poverty has been linked to poorer working memory in adulthood and crucially this relationship seems to be mediated by chronic stress through allostatic load (Evans & Schamberg, 2009). Allostatic load refers to the biological impact of chronic stress often signified by inflammatory responses, activation of the HPA axis and increased pressure on the cardiovascular and metabolic systems (McEwen & Gianaros, 2010). Moreover, the biological effects encompassed in allostatic load have consistently been associated with low SES (Braveman & Gottlieb, 2014; Misiak et al., 2022; Seeman, Epel, Gruenewald, Karlamangla, & McEwen, 2010).

1.2.2 SES and mental health

Low SES has also been associated with increased risk of mental illnesses. Evidence suggests that the association between low SES and increased risk of mental ill-health starts at an early age and persists throughout adulthood (Reiss, 2013). The mental health problems associated with low SES are, amongst others, schizophrenia, ADHD, bipolar disorder, and depression (Agerbo et al., 2015; Gilman, Kawachi, Fitzmaurice, & Buka, 2002; Hakulinen, Musliner, & Agerbo, 2019; Russell, Ford, Williams, & Russell, 2016).

The association between socioeconomic status and depression has been evident for a long time, as mentioned earlier a categorisation of depression from the 17th century argued that poverty could lead to depression (Brink 1979). The association has also been documented in both high, middle and low income countries (Ambugo, 2014). However, the direction of causality is unclear, longitudinal studies of the association are inconsistent and a recent meta-analysis showed associations

between socioeconomic status and depression in only half the studies (Richardson, Westley, Gariépy, Austin, & Nandi, 2015). Recently developed Mendelian randomisation (MR) methods have been used in an attempt to pin down the direction of causality by using genetic instruments of both depression and socioeconomic status as proxies. (Harrison et al., 2020) did find several associations between depression and social factors such as loneliness but also household income. However, they did not find convincing evidence of a direction of causality between SES and depression.

As a common risk factor for both depression and low SES lifestyle factors may explain part of the underlying biological link between the two. As mentioned earlier most research has focused on tobacco smoking, alcohol consumption and obesity. However, there is evidence that while there might be an overlapping risk these lifestyle factors far from explain the whole picture. (Groffen et al., 2013) found that smoking, alcohol consumption, BMI, and physical inactivity was indeed associated with low SES but it did mediate the association between low SES and depressive symptoms. Furthermore, a meta-analysis of MR studies into risk factors associated with depression found no causal association between smoking, alcohol consumption, BMI and depression risk (Köhler et al., 2018). The evidence from MR studies is still inconsistent, a recent MR study into the causal influence of high BMI on depression risk, revealed that a higher BMI was causally associated with an increased risk of depression particularly in women (Tyrrell et al., 2019).

The overlap in diseases comorbid with low SES and depression may be able to hint at a biological mechanism underlying the risk relationship. The overlap in comorbid diseases include cardiovascular disease, diabetes, arthritis, asthma, and obesity. Similarly, the overlap in environmental exposures associated with low SES and environmental risks associated with depression (toxins, air pollution, poor housing) may account for some of the mechanistic association between low SES and depression. Crucially a commonality among the overlapping comorbid diseases and the consequences of the environmental exposures is increased inflammation.

1.3 Inflammation

In short inflammation is a biological mechanism for tissue repair following an injury. During acute inflammation hormones (amongst others, histamines) cause micro vessel dilation thereby increasing blood flow allowing more immune cells (such as lymphocytes and monocytes) to reach the affected area. These immune cells in return increase fluid build-up of the affected part of the body.

Inflammatory reactions can roughly be divided into four groups local vs. systemic and acute vs. chronic. Local inflammation is most often in response to a localised injury or infection and only affects a small part of the body. Local inflammation can still be dangerous, an example of that is that of haemorrhagic stroke. A haemorrhagic stroke is a bleed in the brain which rapidly causes microglia in the brain to release proinflammatory factors causing an acute cascade of inflammatory reactions that can be detrimental to recovery if not treated. Systemic inflammation refers to a body-wide activation of the immune system. Sepsis is a particularly severe and dangerous example of acute systemic inflammation. Often systemic inflammation refers to a low-level chronic inflammation. Chronic inflammation is a slow ongoing inflammation lasting months or years.

Chronic inflammation is associated with diseases like diabetes, cardiovascular disease, arthritis, allergies, COPD and asthma (Okin & Medzhitov, 2012). Chronic inflammatory diseases are relatively common, it is estimated that more than 50% of all deaths are attributable to inflammation related diseases (GBD 2017 Disease and Injury Incidence and Prevalence Collaborators, 2018). In their 2019 review Furman and colleagues point out that the most common causes of systemic chronic inflammation are chronic infections, lack of exercise, obesity, disruption of gut biosis, diet, chronic stress, disturbed sleep and environmental toxins (Furman et al., 2019).

Increased stress can cause disruptions to the balance of internally produced glucocorticoids, that function as anti-inflammatory agents through increased cortisol (a type of glucocorticoid) secretion. This increase in baseline cortisol levels can lead to glucocorticoid receptor resistance thereby disrupting the inhibitory circuit leading to systemic chronic inflammation (Chandola, Brunner, & Marmot, 2006; Cohen et al., 2012).

Treating chronic inflammatory diseases is often difficult as the precise origin of the sustained inflammatory response is unclear and overall an inflammatory response is needed for survival (Tabas & Glass, 2013). One of the most common and effective ways of treating chronic inflammation is with the hormone glucocorticoids (Barnes & Adcock, 2009) thereby supporting the body's own production of the hormones. Glucocorticoids, one of the most common types is cortisol, have wide ranging functions in the immune system where they up-regulate anti-inflammatory proteins and down regulate the expression of proinflammatory proteins. In metabolising glucose, particularly in the liver through gluconeogenesis. Developmentally glucocorticoids are involved in hippocampal development and axonal and dendritic remodelling (Sonia J Lupien, McEwen, Gunnar, & Heim, 2009). As a result, glucocorticoids also play a role in cognitive processes and memory formation (S J Lupien, Maheu, Tu, Fiocco, & Schramek, 2007).

The cascading pathways of chronic stress elevating baseline levels of the glucocorticoid cortisol, thereby disrupting of the HPA axis, causing a sustained inflammatory response and the ultimately leading to physical and mental health problems, i.e. allostatic load, repeatedly been suggested as the explanation for the link between low SES and ill health. However, the mechanistic links between low SES and mental health problems, in particular depression are less well understood.

Systemic chronic inflammation is not only associated with the physical health problems routinely associated with low SES but also the mental health problems associated with low SES.

A longitudinal study of 4,500 children, spanning nine years found that elevated levels of inflammatory markers IL-6 and CRP in blood at age nine were associated with increased risk (OR = 1.55, CI = 1.13-2.14) of clinical depression and more severe depressive symptoms at age 18 (Khandaker, Pearson, Zammit, Lewis, & Jones, 2014). CRP levels in the blood have been found to be elevated in depressed individuals in general but particularly in people with treatment resistant depression (Chamberlain et al., 2019). A meta-analysis of randomized control trials found a lower post-treatment depressive symptom severity in people treated with anti-inflammatory medication than the people treated with a placebo (Husain, Strawbridge, Stokes, & Young, 2017).

Genetic studies of depression have also implicated immune processes in the risk of depression. A GWAS of depression found a considerable number of genes associated with depression in the major histocompatibility complex (MHC) region (Naomi R Wray et al., 2018). The Genes in the MHC region are linked to inflammation regulation, and adaptive as well as innate immune responses (Shiina, Hosomichi, Inoko, & Kulski, 2009). A more recent depression GWAS similarly found genetic variants previously associated with effects of anti-inflammatory- and immunosuppressant drugs (Howard et al. 2019). A recent study on both serum and DNAm derived CRP levels, depressive symptoms and brain structure found a positive association between serum CRP levels and depressive symptoms ($\beta = 0.073$, $p_{FDR} = 0.033$) and a negative association between DNAm CRP levels and global cortical volume ($\beta = -0.1$, $p = 0.01$) (C. Green et al., 2021).

It has been argued that inflammation, as a cause of depression, has an evolutionary role (Miller & Raison, 2016). Following the so-called fight or flight response the body triggers an immune response to the outside threat, often as a consequence of increased activity in HPA axis. The crux of the theory posits that humans have evolved a bias towards inflammatory responses because it protects the host and increases the chance of survival and ultimately reproduction (C L Raison & Miller, 2013). It is argued that inflammatory activity triggers a behavioural response classified as ‘sickness behaviour’ to conserve energy that in turn could be used to facilitate wound healing and fighting infections. Such an evolutionary theory of inflammatory depression may explain why the many genetic alleles associated with depression are so very common in the general population; for the vast majority of human history, they served a genuine purpose.

It is important to point out that evolutionary theories are notoriously difficult to falsify and as such should be treated with some caution. However, it provided a useful framework to explain the observed links between depression and inflammation and why they may have persisted through evolution.

The predominant theory of depression has for a long time been the monoamine theory, focussing on differences in concentrations of neurotransmitters, particularly serotonin, dopamine and norepinephrine, and their association with depression. At first glance it may seem that the inflammatory theory of depression goes against decades of

research and successful anti-depressant treatments. However, they may in fact be closely interlinked and the bridge between them may be histamines.

A recent review has suggested histamines as the missing link between the monoamine hypothesis of depression (dysregulation of neurotransmitters such as serotonin) and the cytokine hypothesis of depression (inflammatory depression). In fact, there is plenty of evidence to suggest that inflammation directly impacts monoamines associated with depression (Charles L Raison et al., 2013). The role of histamines in peripheral inflammation is well known (Haas, Sergeeva, & Selbach, 2008). In the central nervous system (CNS) it is less well understood. Neurons and microglia both produce histamines and microglia can, in turn, be activated by histamines to produce proinflammatory cytokines (H. Dong et al., 2014), which means histamines may play a crucial role in neuroinflammatory processes.

Histamines have been documented to be elevated, both peripherally and in the CNS, in response to stress induced inflammation and more importantly these increased levels of histamines resulted in a reduced neuronal serotonin release by interacting with Histamine H3 receptors (Hersey, Hashemi, & Reagan, 2021). Common SSRIs like escitalopram may even exacerbate this response by off-target effects found to slow the reuptake of histamines in animal models and further unbalancing cerebral histamine levels (Hersey et al. 2021).

1.4 Summary

Depression is the most common mental health disorder, and it is the leading cause of disability worldwide. Depression is clinically heterogeneous and is associated with many comorbid diseases and disorders. Treatment of depression with antidepressant medication or cognitive behavioural therapy is highly effective for most patients. However, around 20% of people with depression suffer from treatment-resistant depression.

There is a vast individual and economic cost of depression. Mood disorders are associated with a six-to-eight-year decrease in life expectancy and between 2005 and 2010 depression was associated with a cost of \$210.5 billion in the United States.

Depression is at least partly caused by genetic factors and the most recent GWAS of depression found 102 independent variants associated with the disorder. Environmental risks for depression include childhood trauma, stressful life events, lifestyle factors (BMI, smoking, and alcohol consumption), environmental toxins, and socioeconomic status. At the intersection between genetic and environmental risk sits DNA methylation.

DNAm levels changes with environmental exposures, these changes are mechanistically specific, and, in some cases, there is evidence of a dose-dependent relationship between environmental exposure and change in DNAm. As a result, DNAm provides a unique opportunity to explore the biological impact of environmental exposures on the individual.

Low SES is a major risk factor for poor health, both physical and mental, and shorter life-span in general. The underlying causes of the low SES and ill-health associations are not well understood but many of them overlap with the risk factors for developing depression. Many of the risk factors associated with both low SES and depression are of an inflammatory nature

1.5 Aims

The association between low SES and depression is not well understood. We lack a cross cultural estimate of the increase in depression risk associated with low SES. There are theoretical frameworks aimed at explaining the association between low SES and depression, but the underlying biological mechanisms are not well understood. The aim of this PhD was to:

1. Establish a cross-cultural and cross-developmental estimate of the risk of depression associated with low SES.
2. Explore the biological impact of living in low SES.
3. Look for possible mechanisms connecting depression and the biological impact of living in low SES.

A systematic review and meta-analysis of all studies, irrespective of country of origin, investigating the association between the risk of depression and low SES was

conducted across countries and developmental time-periods, to quantify the association between the two factors. Next, the plastic and environmentally determined properties of DNA methylation were used to explore the biological impact of living in low SES. To do that an epigenome wide association study (EWAS) of area-based multiple deprivation was conducted. Finally, an EWAS of the interaction of depression and the strongest biological signal from living in low SES, namely inflammation, was performed to look for possible biological links between them.

2 Socioeconomic Status and Depression – A Systematic Review

2.1 Introduction

Major depressive disorder (MDD) is the leading cause of disability world-wide, affecting over 322 million people (WHO 2017). Depression is associated with an elevated mortality risk (Gilman et al., 2017), not least because of the increased risk of suicide in depressed people (Bachmann 2018), but also because of a non-trivial number of diseases comorbid with depression (Walker et al 2018.). While there is a significant genetic contribution to the risk of developing depression (Sullivan et al 2000) there is strong evidence of numerous environmental risk factors, amongst them poverty, deprivation, and low socioeconomic status.

Socioeconomic status (SES) refers to level of income, education and occupation, where lower SES reflects overall material and social deprivation. Lower SES has previously been associated with an increased risk of developing depression (Lorant et al., 2003), decreased life expectancy (Lleras-Muney, 2005), and is one of the strongest predictors of overall poor mental and physical health (Everson, Maty, Lynch, & Kaplan, 2002). Low SES is also risk factor for many of the chronic diseases associated with or comorbid with depression (Hosseinpour et al., 2012; Robbins, Vaccarino, Zhang, & Kasl, 2005).

Despite a growing body of research linking low SES with poor health, this association continues to persist. It is, however, clear that the relationship between health and SES is not static. Globally life expectancy is increasing, education levels are rising and generally the number of people living in poverty is steadily decreasing (Roser, Ortiz-Ospina, & Ritchie, 2013; Roser & Ortiz-Ospina, 2016). However, this global trend is rarely linear or uniform across all countries. For instance, during the Covid-19 pandemic the number of people living in poverty rose for the first time in 20 years (World Bank Group 2020). In a recent government report from the United Kingdom it was concluded that over the last decade the life expectancy of people with

a low SES had in fact decreased and people living in the most deprived areas spent a larger proportion of their life in ill health (Marmot 2020). The report found suboptimal conditions for social mobility suggesting that SES levels are relatively constant across life. The fact that levels of deprivation are improving but an association between SES and depression remains, may be another argument against the threshold theory of SES and health described in the general introduction.

Given the persistent nature of relationships between low SES and health outcomes, such as increased risk of depression, there is a clear need for detailed research on the extant literature to quantify the risk as well as identifying possible differences in the strength of the relationship across subcategories of SES and income economies. Previous research on the association between SES and depression have shown evidence of an association in specific cultures, age groups, and sections of society. The most recent systematic review and meta-analysis on SES and depression by Richardson et al. (2015) found a significant association between neighbourhood socioeconomic conditions and depression in seven out of fourteen studies ($N = 6,833,635$). Their systematic review focussed exclusively on longitudinal studies and sensitivity analyses revealed that results varied by length of follow-up. Studies with less than five years of follow-up found higher odds of depression with lower SES, but studies with five years or more of follow up showed no association between SES and depression. A finding that may be in part attributable to biased attrition.

The studies included in Richardson et al. (2015) were all from high-income countries (USA, Canada, Australia, United Kingdom, and Sweden) and all, but two studies were from English speaking countries. Half the studies included in the review were restricted to people living in urban areas. The vast majority of the participants (99.5% of the total sample) in the review come from two studies (Lofors & Sundquist, 2007; Sariaslan et al., 2015) both of which are from Sweden. One of those studies (Lofors et al. 2007) is a population-wide register-based study including all adults aged 25-67 years old in 1997, 4.5 million in total. The other study included all children born 1967-1989, 2.3 million in total (Sariaslan et al., 2015). That means a large proportion of the newest study was also included in the former and thereby counted twice. Assuming a steady rate of births, the estimated overlap is about 25% of the Sariaslan et al. (2015) study sample, i.e. over half a million participants. It is also worth pointing out that the studies included in the Richardson et al. (2015)

meta-analysis have a spectrum of depression measures ranging from sub-clinical depressive symptoms to hospital admissions due to major depressive disorder. The inclusion of depression measures of widely ranging severity decreases the specificity of the meta-analysis results to clinical depression.

The aim of the Richardson et al. (2015) systematic review and meta-analysis was to investigate the influence of neighbourhood factors instead of individual level factors. This approach offers an insight into the geographic- and population level patterns of depression but it does not allow for more detailed inferences about circumstances of the individual and how they relate to depression. Furthermore, the review was restricted to studies published in peer reviewed journals. Whilst this may reduce the risk of including low quality studies, it also increases the risk of introducing publication bias into the meta-analysis. These relatively narrow inclusion criteria of Richardson et al. (2015) facilitate the secondary aim of their systematic review namely, to explore whether methodological factors can explain heterogeneity in previous findings. However, it also hinders the generalisability of the finds. As a result, it is possible that the findings, based on 14 studies from high-income, western, and predominantly English-speaking countries, may not apply to countries and cultures outside of this group.

A systematic review by (Lund et al., 2010) focused exclusively on low- and middle-income countries when investigating the association between poverty and common mental disorders (CMD), including depression. The systematic review included 115 studies, 71 of which focussed on depression with a further 30 studies that focussed on other CMD. The majority of included studies showed a significant association between poverty measures and CMD. A more detailed inspection of the poverty measures revealed that SES overall, education, social class, housing, financial stress, and food insecurity were overwhelmingly associated with CMD, whereas the association between CMD and income, employment, or consumption was less clear.

This review is an important addition to the literature on the association between SES and depression by providing an overview of studies from low- and middle-income countries. Similarly, the larger number of included studies from both clinical and community settings increase generalisability and helps to minimise inflation of prevalence. However, the studies were not meta-analysed so a pooled estimate of the size of the association was not available. The population studied is adult and as a result

may miss developmental periods in which the association between SES and depression differs in magnitude. Finally, any inferences drawn from the study are not specific to depression but about CMD, including anxiety and somatoform disorders.

The association between SES and depression in children was investigated by (Reiss, 2013) in their systematic review on SES inequalities and mental health in children and adolescents. Most of the 55 included papers, found an association between SES and mental health in general. Some of the included studies found a significant negative association between depression and SES in children. One of the longitudinal studies included in the Reiss (2013) review showed an effect for anxiety but not depression. However, as the review (Reiss et al. 2013) was focussed on broad mental health there were no inferences made about the association between SES and depression. The systematic review did not perform a meta-analysis, so an estimate of a pooled effect size was not provided. The inclusion criteria restricted the publication date of the studies to be between 1990 and 2011 and finally, the majority of included studies are from North America, Europe and Australia, i.e., high income countries.

Previous systematic reviews into the association between SES and depression have made important contributions to defining and quantifying the association in several subgroups of populations. However, the current literature is missing an up-to-date investigation into the existence and size of an association between SES and depression across cultures and age groups. This systematic review and meta-analysis seek to provide a global estimate across cultures and participant ages, conducting meta-analysis where possible to derive quantitative estimates of risk sizes.

Throughout the last few decades there have been increases in standard of living in many low- and middle-income countries (Permanyer & Smits, 2020) while some countries have experienced a slowing in socioeconomic development and even increases in inequality including health inequality (Marmot report). These changes to poverty, inequality, and SES in different directions for different countries adds to the importance of an up-to-date estimate of the association between SES and depression.

We need to understand the strength of the relationship between SES and depression to aid changes to policy that can accommodate preventative strategies around depression. This understanding must be based on the largest possible body of empirical research, and it has to be based on evidence from all indicators of SES, be it

education, income, occupation, neighbourhood or individual level factors, and finally it has to represent and be generalisable to diverse cultures and geographies.

Given the above highlighted gaps in existing literature the following systematic review and meta-analysis aim to examine associations between SES and depression. Studies available in the English language were included in the investigation of the association between clinically relevant depression and socioeconomic status, without any limits on participant age, publication date and country of origin. The aim of the present chapter is to meta-analyse the effect of SES on the risk of developing depression both for single indicator factors and composite measures of SES. This systematic review aims to include studies from all cultures and economic backgrounds. Finally, studies will be included and meta analysed regardless of whether they are done in a clinical or community setting and whether they are of a longitudinal or cross-sectional design.

Studies on perinatal depression were excluded from the review and metanalysis. Perinatal depression is circumscribed to a specific life stage and may have different antecedents and causes to other forms of depression. Therefore studies of perinatal depression were excluded from this review.

2.2 Methods

2.2.1 Inclusion and exclusion criteria

In the present review peer-reviewed literature and published PhD theses that examine the association between SES and clinically relevant depression have been included. PhD theses were included to mitigate publication bias. Furthermore, for a number of the included studies the investigation of the association between SES and MDD is not the primary outcome. As a result, several studies report null findings on the association between SES and depression, suggesting that risk of publication bias may be reduced or mitigated. Studies that reported findings on the distribution of participants with or without clinically relevant depression across SES as a composite measure or any and all of the sub parts of the measure, i.e. income, education and employment, were included. Studies investigating depressive symptoms or subtypes of depressive disorders were not included in this review. There were no restrictions on the type of study (longitudinal or cross sectional), the country of the study population, the setting of the study (clinical or community) or the date of publication.

Studies involving participants with clinical depression as diagnosed with any recognised diagnostic criteria, whether self-reported, obtained through data linkage, or diagnosed using a diagnostic clinical interview or questionnaire with a clear clinical cut-off, were included. Studies focusing on symptoms and comparisons between subgroups of depression were excluded. Studies with an unclear division between cases and controls i.e. comparing sub groups of depressions, were excluded. It is not current authors prerogative to make decisions on behalf of the included study authors as to which subgroup of depression fits with cases and which subgroup fits with controls. Rather it is our task to report on, and aggregate existing findings.

There were no restrictions on the age, sex, cultural heritage or comorbid diseases of the participants in the present review.

Studies that reported insufficient results to be meta-analysed were excluded from the review unless the study authors provided the necessary data. For example, studies that failed to report number of participants overall, number of cases or number of controls. In other instances, analysis was done on a non-transparent weighted basis that hindered meta-analysis. In all these occasions, study authors were contacted to request data in a raw-form or in a form conducive to meta-analysis. If this data could not be obtained the study was excluded. All authors were contacted by email, except in two cases, one where the corresponding author was deceased and one where the corresponding author's contact details with the military was no longer in use. Out of 54 studies 18 authors replied and one author was able to provide sufficient additional data.

Studies that have grouped depression with other psychiatric disorders (often anxiety or schizophrenia) in the "case" group and compared it to a "control" group without both the psychiatric disorders present in the "cases" group have been excluded. However, there were no restrictions on comorbidities being included in the present systematic review. As a result, studies with participants with more than one psychiatric disorder or health problem were included, as long as the only difference between the cases-group and controls-group was the presence or absence of depression.

2.2.2 Search strategy

Literature searches were performed in the following databases: PsycInfo, EMBASE, Medline, Web of Science, ASSIA. The databases were chosen to cover as broad a range of medical and sociological sources as possible.

In four out of five databases searches were conducted using subject headings. The following search template was used in PsycInfo, EMBASE and ASSIA: Depression AND (Socioeconomic status OR Income OR Occupation OR Education).

Web of Science does not facilitate subject heading searches. Instead, a keyword search was conducted using the same key words that were used in the subject heading search (Depression ADJACENT (Socioeconomic status OR Income OR Occupation OR Education)) ensuring the keywords were within 5 words of each other using the “adjacent” search function.

In all databases filters were applied to exclude reviews, conference abstracts, animal research and limit the search to English language papers. There were no restrictions on publication date or country of origin of the search results. Studies on perinatal depression, depression subtypes and depressive symptoms were not filtered out. This was done to minimise automatic exclusion of studies meeting inclusion criteria. Instead, studies investigating perinatal depression, depression subtypes and depressive symptoms were excluded during the two screening processes.

The last date for all searches was 15th of March 2019.

2.2.3 Data collection

Searches from the 5 databases returned 8312 studies. 369 duplicates were removed using Endnote software. 7943 studies were screened against their title and abstract, 7422 studies did not meet inclusion criteria and were excluded. Studies screened against their title and abstract, that met inclusion criteria but could not be found thorough online databases were requested through the University of Edinburgh’s interlibrary loan services. Assessing interrater reliability is important to ensure that the screening process does not miss any eligible studies and that the screening method is replicable across different assessors. Selection bias can be introduced in the screening processes if an assessor repeatedly includes or excludes

studies based on factors other than those listed in the inclusion and exclusion criteria. To assess interrater reliability and selection bias a second author (RM) screened a 10% subsample (N = 795) of abstracts and titles. There was an 89.7% agreement on the outcome of the screening by the two authors. 521 studies were screened against their full text.

Data collection was done alongside full text screening. The following data was collected on papers found to meet inclusion criteria: author(s), year of publication, journal of publication, name of country data is from, economic category of country, which cohort the data is from, the setting of data collection (community or clinical), study design (cross sectional or longitudinal), if longitudinal the number of years of follow-up, overall sample size, the number of controls, the number of cases, the proportion or number of female participants, the age of participants, Groups (case control), Depression diagnostic tool, comorbidities, number of SES groups, statistical estimates for SES, Income, Occupation or Education, variables adjusted for, and comments.

In studies reporting dichotomous or ordinal data for the SES, income, occupation, and education variables, odds ratios (ORs) or risk ratios (RRs) and confidence intervals, or raw case-control data were collected for each SES group. For studies reporting continuous results for the SES, income, and education variables, mean differences and standard deviations were collected. Some studies reported raw data in the form of percentages for which the prevalence numbers in each of the SES groups were imputed and rounded to the nearest whole number.

Risk of bias refers to the possible inclusion of studies with methodological problems that could raise questions about the validity of their findings and ultimately affect the effect estimates of the meta-analysis. Widely used tools are available to assess risk of bias in clinical trials. However, to date there is no universally recognised screening tool for risk of bias for observational studies (Bero et al., 2018). Since the outcomes of interest in the present review are purely prevalence and observations based, a risk of bias assessment was not used in the screening process. Instead sampling methods of each study was assessed as part of the screening process. No studies were found to involve biased sampling methods.

Non-reporting bias refers to the introduction of bias to the effect estimates of the meta-analysis due to missing results. The most common cause of missing results is due to biases in publications of results based on effect sizes, direction of effect, or even p-value of an inferential statistic. Non-reporting bias is not considered as a big problem in the present review since several studies have reported null findings of the association between SES factors and depression. One could argue that because the SES-depression association is not necessarily the main outcome of a large proportion of the included studies, and therefore less likely to be the focus of the resulting manuscript it is less likely that those studies have suffered from reporting bias based on the size or direction of the effect.

2.2.4 Analysis

For the meta-analyses the studies were divided into four groups; dichotomous, ordinal with three groups, continuous, and longitudinal. The dichotomous group is made up of all the studies reporting data in a binary/dichotomous format and all the studies reporting results in an ordinal format with 4 groups. The latter type of study was dichotomised by merging the top 2 groups together and the bottom two groups together. In the dichotomous group SES, income, education, and employment were meta-analysed separately. As a result of meta-analysing subcategories of SES separately, some studies appear in several meta-analyses, because they have reported results in more than one SES category.

A pooled OR and 95% CI was calculated for each subgroup using the package ‘meta’ (Balduzzi, Rücker, & Schwarzer, 2019) in R version 4.0.4. The effects were calculated using a random effects model and the between studies variation was calculated using the Sidik-Jonkman estimator (IntHout, Ioannidis, & Borm, 2014). RRs and ORs were meta-analysed together as they are statistically comparable when the prevalence is low (<10%).

Studies reporting results in an ordinal format with three groups were meta-analysed in a sequential manner. An OR estimate and 95% CI was calculated for the low SES vs medium SES comparison and another OR estimate and 95% CI for the low SES vs. high SES. Calculations of ORs and 95% CI was performed in the same way as for dichotomous data. Continuous data was meta-analysed by calculating

standardized mean differences for each SES subgroup with a random effects model. Between study variation was calculated using the restricted maximum likelihood estimator.

A narrative approach was adopted for papers that could not be meta-analysed due to the design of the study or the way in which the data was reported. This was done to make sure there was no systematic exclusion of papers due to differences in methodological approaches. Any study with the SES group divided into five or more categories were included in the narrative review. In instances where occupation was divided into more than two groups (employment, unemployment) the measure was not meta-analysed but instead included in the narrative review. This was done because the incremental distance between each occupation group other than employed versus unemployed (i.e. student, retired, full-time, part-time, manual, non-manual, public private, etc) is not consistent. Furthermore, different occupational groups do not translate well across studies let alone cultures.

If a study is represented in any of the meta-analyses it will not be included in the narrative review. The narrative review has been included to make sure that studies are not systematically excluded because of methodological approaches, if however, the study has already been included it is not necessary to include any other SES measures in the narrative review.

All longitudinal papers were included in a longitudinal narrative review due to a combination of low number of studies in each of the SES groups and variable follow-up time.

Significant heterogeneity between studies was anticipated given the larger number of studies included and the broad inclusion criteria with regard to country of origin, age groups investigated, and diagnostic tools employed. Statistical between study heterogeneity was explored using Cochrane's Q and quantified by an I^2 statistic. Both measures are calculated using the 'meta' package in R.

To explore possible origins of between study heterogeneity, sensitivity analyses based on the economic situation in the country of origin were performed. Using the World Bank country and lending group classifications I divided countries into low- and high-income economies. A sensitivity analysis was carried out for each

meta-analysis on dichotomised data for a particular SES group. The sensitivity analysis compared ORs and CIs and heterogeneity measures for each SES measure income economy subgroup. Table 2.1 shows the Gross National Income (GNI) per capita brackets used by The World Bank to classify a country's income economy. The GNI brackets are based on the 2022 World Bank classifications despite all included studies being published prior to March 2019. For the purpose of the sensitivity analyses it is assumed that whilst the majority of countries have seen an increase in GNI, the relative distribution of countries into high vs. low-income groups has seen negligible change.

World Bank Country Group	GNI per capita (\$)
Low-income economy	0-1045
Lower middle-income economy	1046-4095
Upper middle-income economy	4096-12695
High-income economy	12696-

Table 2.1: Gross National Income (GNI) per capita brackets for each income economy group

2.3 Results

2.3.1 Study inclusion

The literature search returned 8,312 studies, 369 of which were duplications. During title and abstract screening of the 7,943 remaining studies, 7,422 were excluded. The remaining 521 studies were full text screened, leading to another 381 papers being excluded. The majority of these papers (188) were excluded because the study assessed symptom severity rather than clinical depression, The second biggest group were excluded because there was no SES MDD comparison (78), insufficient reporting of results accounted for 60 excluded studies, 17 studies were excluded because MDD had been grouped with another psychiatric disorder in the experimental group, 10 excluded studies were not available in English, 4 studies could not be found. The data from 4 studies had already been published in other included studies, another 4 studies did not include a comparison with a control group, 4 studies were on perinatal depression, another 3 were reviews, 2 studies were excluded because they

were duplications, and another 4 references were book chapters, comments, or conference proceedings. Finally, a study was excluded because it had matched cases and controls for SES, one study was qualitative and finally one study was excluded for quality reasons; the reported data and results did not match up.

As a result, 140 studies were included in the systematic review and meta-analysis. Longitudinal studies (15) were included in a narrative review as low study numbers for each SES category and differences in follow-up time did not allow for a meta-analysis. Similarly, studies reporting results in an ordinal format with more groups than 4 and studies that could not be meta-analysed because there was only one study in that category were also included in the narrative review (N = 13).

The meta-analysis therefore consists of 112 studies, with analyses divided into three major groups; a dichotomised data analysis, analysis of data reported in 3 SES groups, and a continuous data analysis. A dichotomised meta-analysis was performed for each SES measure; SES, income, education and employment. The ordinal meta-analyses of data reported in 3 SES categories were performed for each SES measure; SES, income, and education. Employment was not included in the ordinal meta-analyses as it did not make methodological sense to meta-analyse employment status over three categories. Continuous data were available for meta-analysis for income and educational attainment. Below is a flowchart (Figure 2.1) with the number of studies excluded at each step and the number of included studies in each of the meta-analyses.

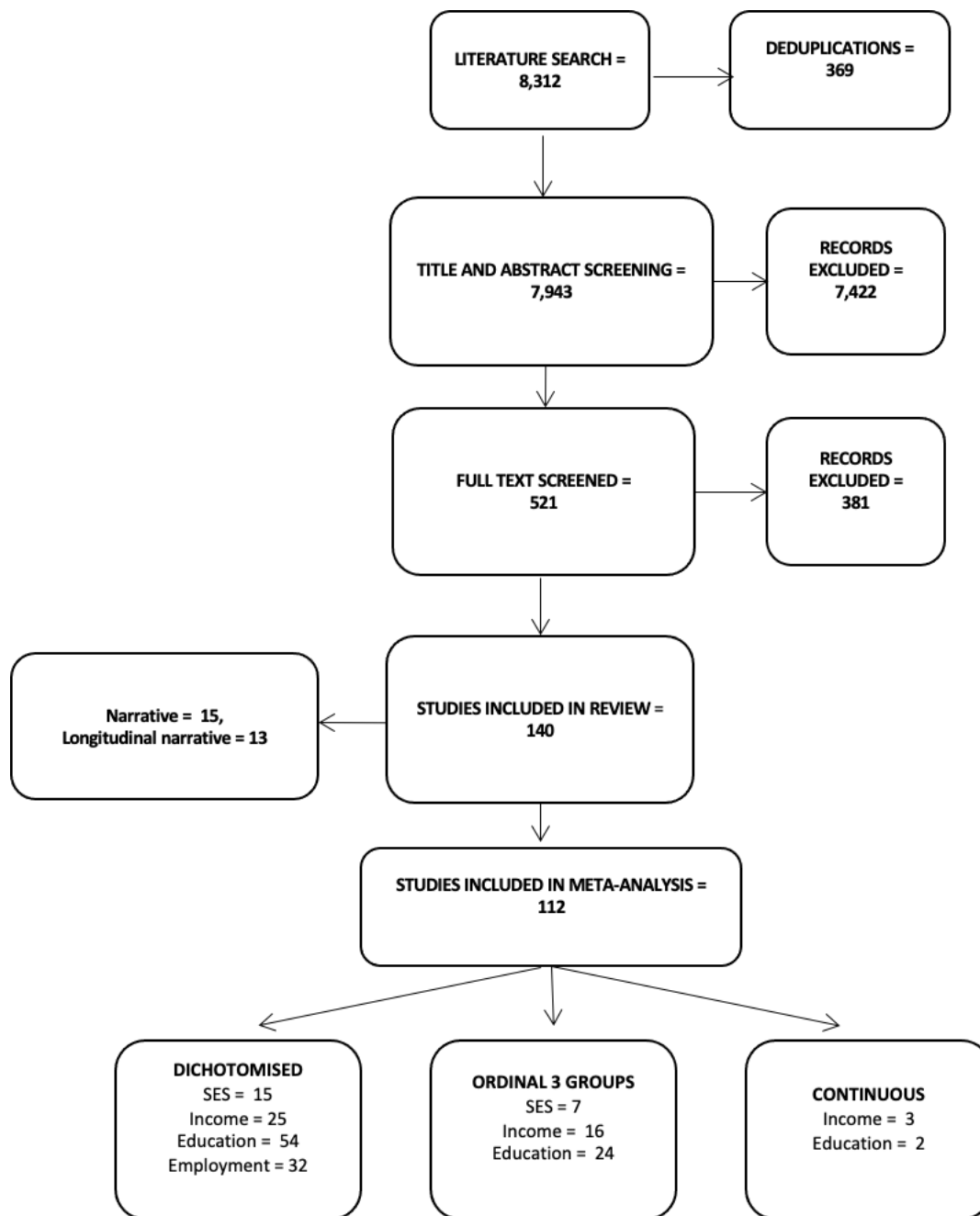


Figure 2.1: Flowchart of number of included and excluded studies. Studies in the “Dichotomised” bin reported data in 2 groups (e.g. high vs. low income), studies in the “ordinal 3 groups” bin reported data in 3 groups (e.g. low, medium, and high educational attainment), and studies in the “continuous” bin reported continuous results (e.g. median income for depressed people).

2.3.2 Missing data and contact with authors

Authors from 57 out of the 60 studies that reported insufficient data to allow meta-analysis, were contacted. Author contact details for two studies could not be

found and another had since deceased. Out of the 57 authors that were contacted, 18 replied, and 1 provided supplementary data enabling the study to be included in the meta-analysis.

2.3.3 Study population

The total number of participants in the 140 studies is 329,005 of which 48,158 presented with depression. The resulting overall prevalence is 14.6%.

The participants were recruited from 47 countries across 6 continents. Below is a map (Figure 2.2) of the geographic distribution of depression cases from the meta-analysis. The map illustrates the inclusion of cases from across the world, although USA, Canada, Brazil, the United Kingdom, France, South Korea, and Japan are strongly represented while the number of African countries represented and the number of cases from these is comparatively low.

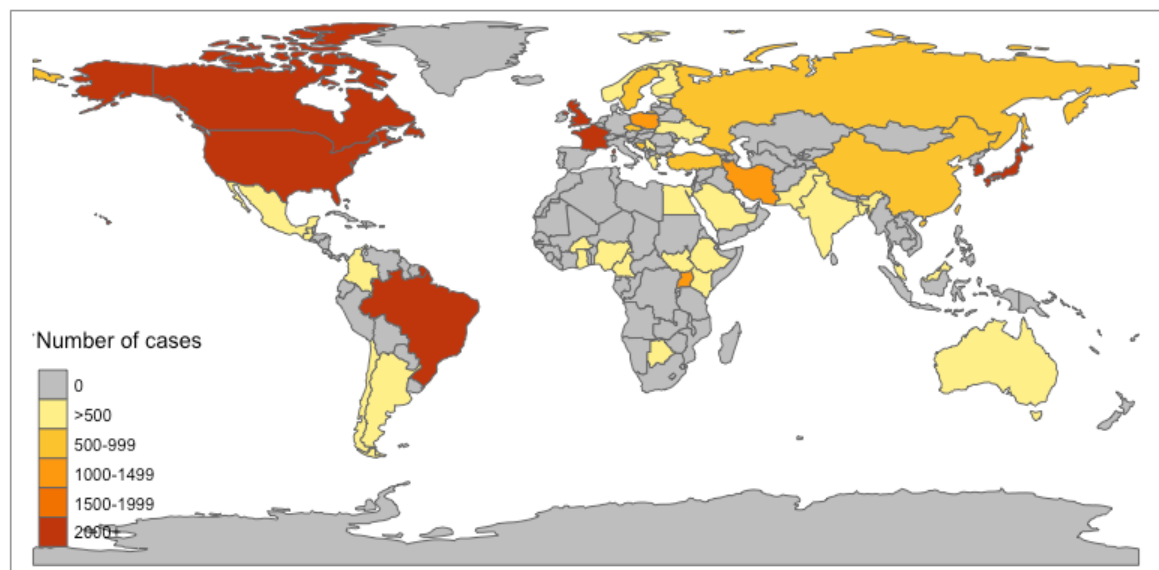


Figure 2.2: Map of distribution of cases

Of the studies included in the meta-analysis, 63 are from high income countries, 30 from higher middle-income economies, 11 from lower middle-income economies and finally 10 from low-income economies. The classification of low-, middle- and high-income countries is based on the World Bank definition (The World Bank 2022). These numbers illustrate a better economic spread of the countries

included in the meta-analysis compared to previous literature. However, there is still a skew toward high and higher middle-income countries.

Most studies included middle aged participants. In 56 of the studies reporting statistics on their participants age, participants were over 18 and had a mean age below 50. Children were included as participants in 10 of the studies and in 31 studies participants had a mean age of over 50.

Finally, of the studies included in the meta-analysis 72 of them recruited in a community setting while 39 of them recruited in a clinical setting.

2.3.4 Dichotomised meta-analyses

2.3.4.1 Socioeconomic status

87 included studies reported data on SES in a dichotomous manner or dichotomised (joining the top two and bottom two SES groups) to be included in a meta-analysis.

Results from 15 studies were included in the dichotomised meta-analysis of the association between SES and depression. Figure 2.3 is a forest plot of the individual ORs and 95% CIs of each study as well as a pooled result. The pooled OR and CI, calculated using a random effects model, for SES and depression is 1.70 (1.32-2.20), representing a 70% increased risk of depression in individuals of low SES compared to high SES. A null result was found in five of the included studies and no studies found a reverse relationship between SES and depression.

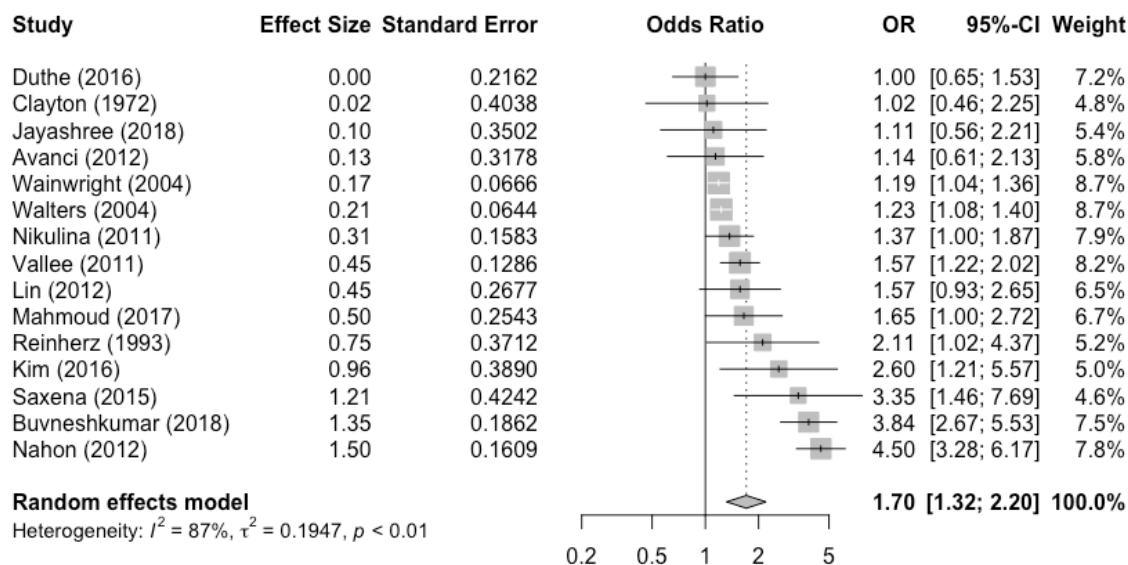


Figure 2.3: Forrest plot of ORs and CIs of depression risk in people with low SES, from dichotomised data. (e.g. high vs. low SES)

The between-study heterogeneity was estimated at $I^2 = 87\%$ (with a variance of $\tau^2 0.19$). Cochran's Q was estimated at 105 with a p value of less than 0.0001.

To investigate the between study heterogeneity a sensitivity analysis of the association between SES and depression was performed, dividing studies into high- and low-income economy groups.

Economy	N of studies	OR	95% CI	Q	I^2	p_{subgroup}
High	10	1.63	1.22-2.18	68.24	86.8%	0.670
Low	5	1.86	1.08-3.22	27.42	85.4%	

Table 2.2: Sensitivity analysis of SES and depression association in dichotomised data.

The results of the sensitivity analysis are presented in Table 2.2. A chi-squared test showed no significant subgroup differences in the estimates of risk ($df=1$, $p = 0.670$) and the I^2 of both subgroup meta-analyses remains high. These results suggest that the heterogeneity is not introduced due to economic differences of the countries from which the studies originate.

2.3.4.2 Income

The dichotomised meta-analysis of the association between income and depression included data from 25 studies. The forest plot below (Figure 2.4) illustrates the effect (OR and CI) of each individual study as well as the pooled effect estimate calculated using a random effects model. The meta-analysis reveals that people on a low income has a 95% increase in odds of having depression compared to people on a high income (OR = 1.95, CI = 1.51-2.53). An inspection of the forest plot reveals that none of the studies found a reverse relationship between low income and depression and nine out of the 27 studies reported a null finding.

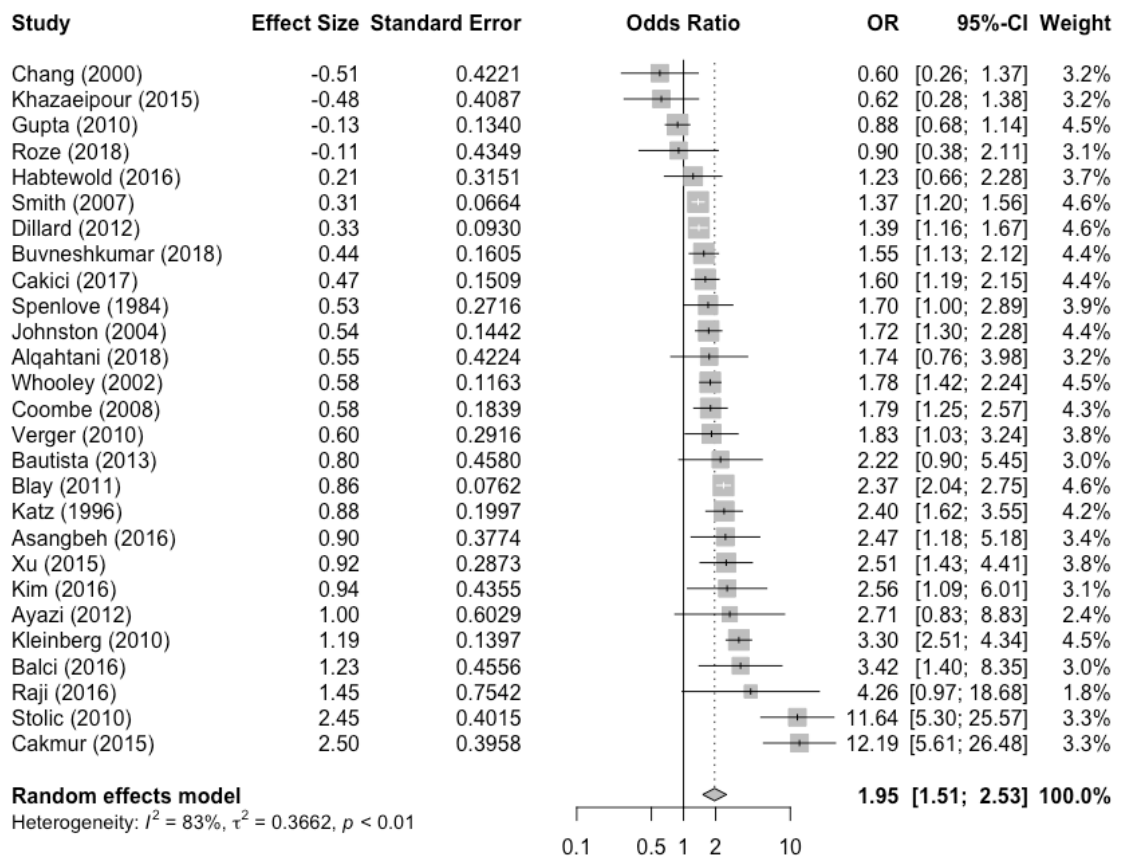


Figure 2.4: Forrest plot of ORs and CIs of depression risk in people with low vs high income, from dichotomised data.

The between-study heterogeneity was estimated at $I^2 = 83\%$ (with a variance of $\tau^2 0.37$). Cochrane's Q was estimated at 156 with a p value of less than 0.0001.

To investigate the heterogeneity, A sensitivity analysis was performed, of the association between income and depression dividing studies into high- and low-income economy groups.

Economy	N of studies	OR	95% CI	Q	I ²	p _{subgroup}
High	22	1.95	1.44-2.63	151.6	86.1%	0.815
Low	5	1.83	1.19-2.82	4.42	9.5%	

Table 2.3: Sensitivity analysis of income and depression association in dichotomised data.

The results of the sensitivity analysis can be seen in Table 2.3. There was an uneven distribution of studies across country economies with 5 studies from a low-income country and 22 from a high-income country. Despite the difference in the size of the confidence intervals a chi-squared test showed no significant difference between subgroups in the estimate of risk ($df = 1$, $p = 0.815$). However, the between study heterogeneity did differ across the two subgroups. The I^2 of the high-income economy group remained comparable to the overall I^2 (86.1% vs. 83%) while the I^2 of the low-income economy group is a negligible 9.5%. This sensitivity analyses suggest the source of the between study heterogeneity is to be found in the group of studies from high income economies. Whether this is caused by methodological differences, cultural differences or simply a difference in the number of studies in each subgroup remains unclear.

2.3.4.3 Educational attainment

The association between educational attainment and depression was assessed using results from 54 studies. A forest plot below (Figure 2.5) presents the ORs and CIs for each individual study as well as the overall effect estimate calculated using a random effects model.

Based on the meta-analysis of dichotomised data, people with a lower educational attainment have a 69% higher risk of having depression compared to people with a higher educational attainment (OR = 1.69, CI = 1.43-2.00). A visual inspection of the ORs and CIs for each individual study reveals that 31 studies found an increased risk of depression in people with low educational attainment, 22 studies

revealed a null result, and one study found a reverse relationship; that high educational attainment is associated with an increased risk of having depression.

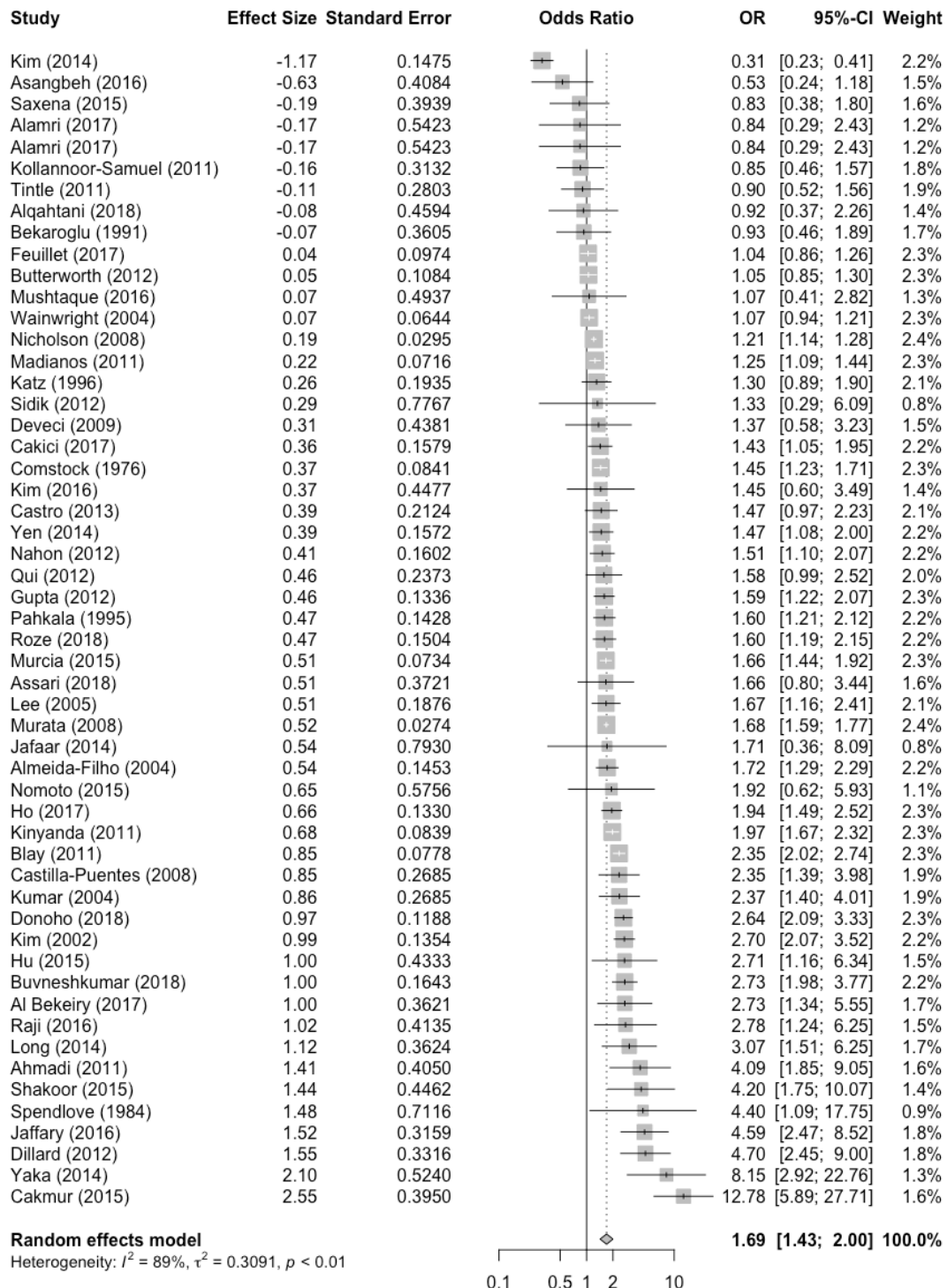


Figure 2.5: Forrest plot of ORs and CIs of depression risk in people with low educational attainment, from dichotomised data.

The between-study heterogeneity was estimated at $I^2 = 89\%$ (with a variance of $\tau^2 0.31$). Cochrane's Q was estimated at 465 with a p value of less than 0.0001.

To investigate whether the between study heterogeneity can be explained by differences between low- and high-income economies, a sensitivity of the association between educational attainment and depression dividing studies into high- and low-income economy groups, analysis was performed.

Economy	N of studies	OR	95% CI	Q	I^2	p_{subgroup}
High	45	1.67	1.40-2.00	423.49	89.2%	0.895
Low	9	1.73	1.05-2.86	39.18	79.6%	

Table 2.4: Sensitivity analysis of educational attainment and depression association in dichotomised data.

The results of the sensitivity analysis can be seen in Table 2.4. A chi-squared test showed no significant difference between risk estimates of the two subgroups ($df = 1$, $p = 0.895$). The heterogeneity estimate, I^2 , does not differ markedly between studies. The results suggest that the between study heterogeneity in the meta-analysis of the association between educational attainment and depression is not due to differences between countries from low- and high-income economies.

2.3.4.4 Unemployment

The meta-analysis on the association between unemployment and depression included dichotomised results from 29 studies. The ORs and CIs for each individual study as well as the pooled OR and CI ($OR = 1.57$, $CI = 1.36-1.80$) from a random effects model can be seen in the forest plot below (Figure 2.6).

The result suggests that people who are unemployed have a 57% higher risk of experiencing depression. A visual scan of the risk estimates from the individual studies reveals that 17 studies found a significant increased risk for depression associated with unemployment whilst, 10 studies reported a null result.

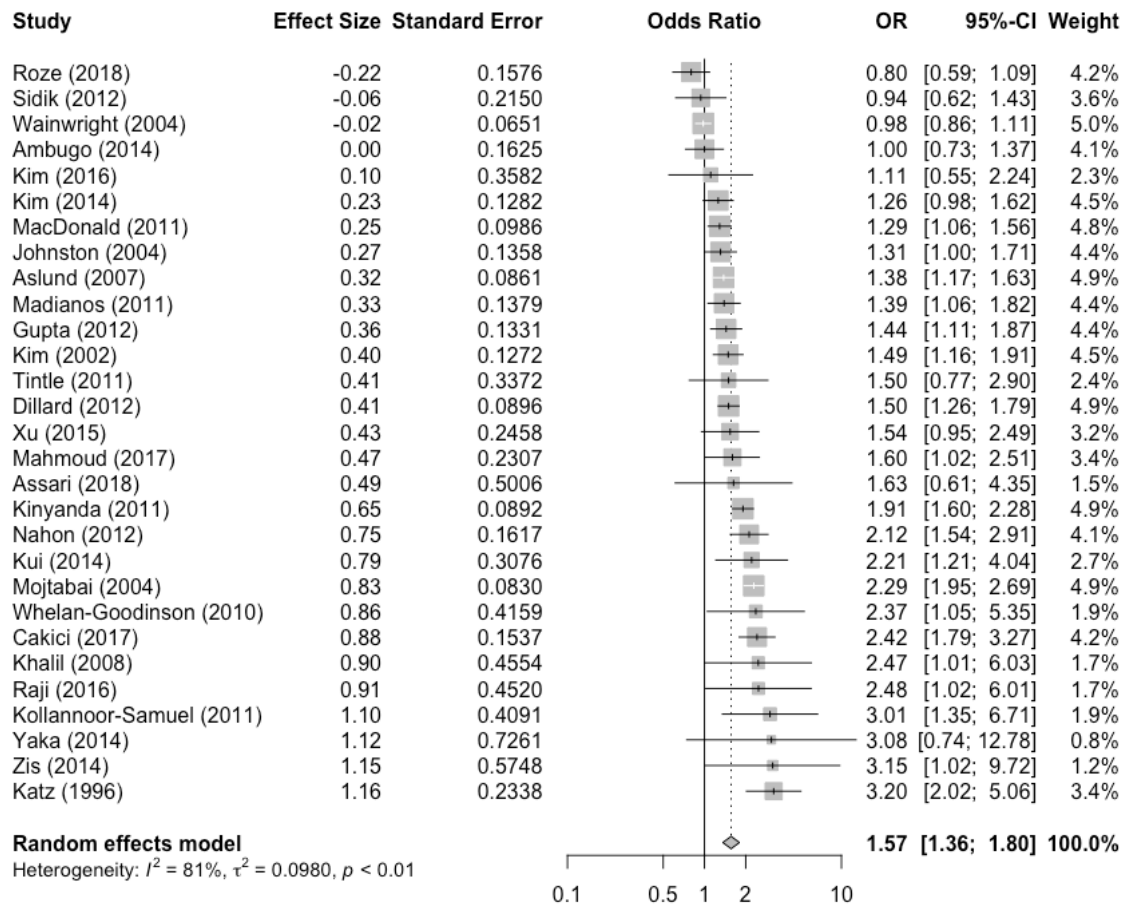


Figure 2.6: Forrest plot of ORs and CIs of depression risk in people with unemployment, from dichotomised data.

The between study heterogeneity was estimated at $I^2 = 81\%$ (with a variance of $\tau^2 = 0.09$). Cochran's Q was estimated at 138 with a p value of less than 0.0001. To investigate the origin of the heterogeneity a sensitivity analysis was performed.

Economy	N of studies	OR	95% CI	Q	I^2	p_{subgroup}
High	23	1.55	1.32-1.83	126.29	82.6%	0.208
Low	5	1.85	1.49-2.31	1.70	0.0%	

Table 2.5: Sensitivity analysis of educational attainment and depression association in dichotomised data

The results of the sensitivity analysis can be seen in Table 2.5. A chi-squared test no significant difference between the risk estimates from low-income countries and high-income countries ($df = 1$, $p = 0.208$).

However, the estimate of heterogeneity differs between the two subgroups. The I^2 for the low-income economy studies is 0% while the high-income economy studies have an I^2 of 82.6% analysis. While the result is not significantly different, the sensitivity analysis not only shows that the association between unemployment and increased risk of depression is more consistent across studies, but the pooled estimate also suggests that the risk of depression in unemployed individuals from low-income economy countries is higher than that of individuals from high income economies.

2.3.4.5 Publication bias in dichotomised studies

To assess the level of publication bias in the included dichotomised studies a funnel plot has been included (Figure 2.7). A visual examination of the funnel plot shows a slight asymmetry suggesting the presence of publication bias. To formally test for asymmetry an Egger's test was carried out which again suggest a significant degree of asymmetry ($z = 2.0551$, $p = 0.039$).

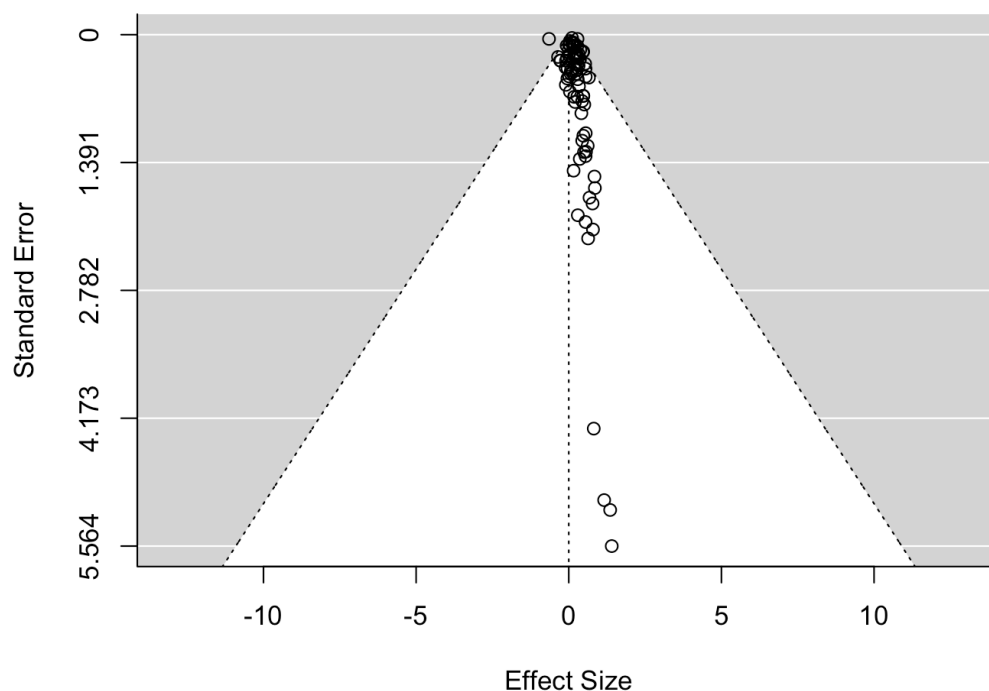


Figure 2.7: Funnel plot of all studies included in dichotomised meta analyses. The y axis represents standard error and the x axis represents effect size. The dashed lines and the white background represents the pseudo 95% confidence interval and each small circle represents one study.

2.3.5 Ordinal three group meta-analyses

Studies reporting results dividing SES measurement into three groups (i.e. low, middle, and high) were meta analysed by comparing the risk of depression in low SES groups with the risk of depression in high SES groups, and comparing the risk of depression in medium SES groups with the risk of depression in high SES groups. This type of meta-analysis creates two ORs and CIs for each SES measurement. Employment is not included in this analysis as it was deemed untenable to group employment into 3 groups in a methodological sensible way across studies.

2.3.5.1 Socioeconomic status

There were seven studies reporting results on the association between depression and SES, using three SES groups. The results of the meta-analysis assessing risk of depression in low versus high SES (OR = 2.20, CI = 1.14-4.24, $p = 0.018$) can be seen in Figure 2.8 and the result of the medium versus high SES (OR = 1.40, CI = 0.84-2.33, $p = 0.195$) can be seen in Figure 2.9.

Based on these seven studies people with a low SES compared to people with a high SES, have a 120% increased risk of having depression. However, it is worth noting that the CI is relatively broad for this risk estimate and a visual inspection of the risk estimates of the individual studies show that four out of the seven included studies found a null effect. There was no significant difference in risk of depression for people with a medium SES compared to people with a high SES, in fact one of the individual studies found a significantly higher risk of depression in people with a medium SES compared to people with a high SES.

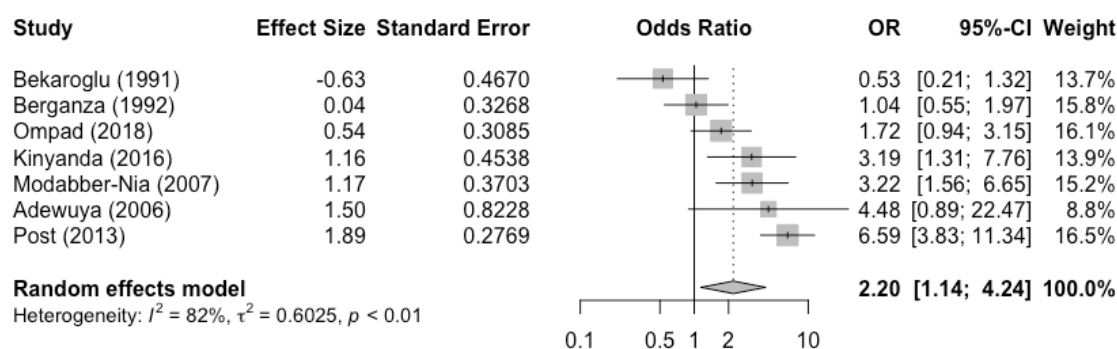


Figure 2.8: Forest plot of depression risk associated with low versus high SES.

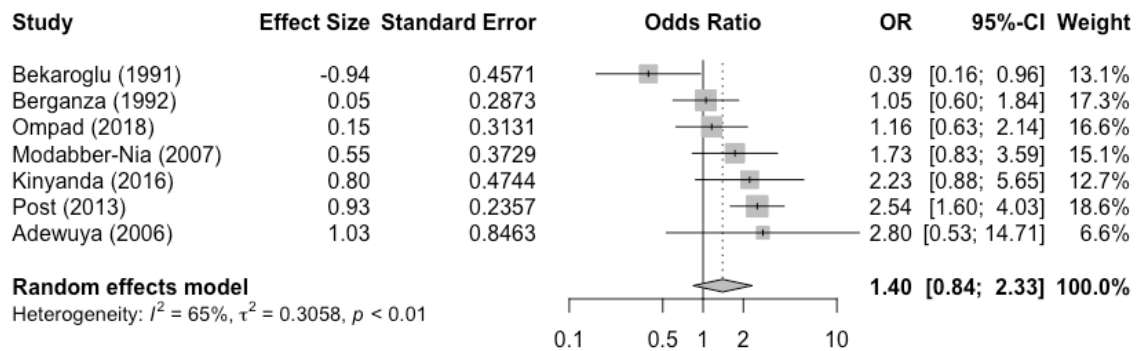


Figure 2.9: Forest plot of depression risk associated with medium versus high SES.

The between study heterogeneity in the low vs. high SES meta-analysis is estimated at $I^2 = 82\%$ with a variance of $\tau^2 = 0.6$. This is a relatively high level of heterogeneity which may reflect the inconsistencies in results of the individual studies. The between study heterogeneity in the medium vs. high SES meta-analysis is comparatively low, I^2 was estimated at 65% with a variance of $\tau^2 = 0.3$. Similarly, the relatively low level of heterogeneity may reflect the consistency of the individual study risk estimates, five out of seven studies found no significant difference in risk of depression in people with a medium SES compared to people with a high SES.

2.3.5.2 Income

The meta-analysis of studies comparing levels of depression across three income groups included 16 studies. The meta-analysis of depression risk in individuals with a low income compared to individuals with a high income revealed a 160% higher risk of depression compared to people on a high income (OR = 2.60, CI = 1.72-3.92, $p < 0.0001$). The risk estimates of the individual studies and the pooled estimate calculated using a random effects model can be seen in Figure 2.10.

A visual inspection of the forest plot of low versus high income meta-analysis show that 11 out of the 16 studies found a significant increased risk of depression in people with a low income compared to people with a high income. Although, the ORs and CIs of the individual studies vary greatly. Five studies reported a null result, and no studies found a reverse association between depression and income.

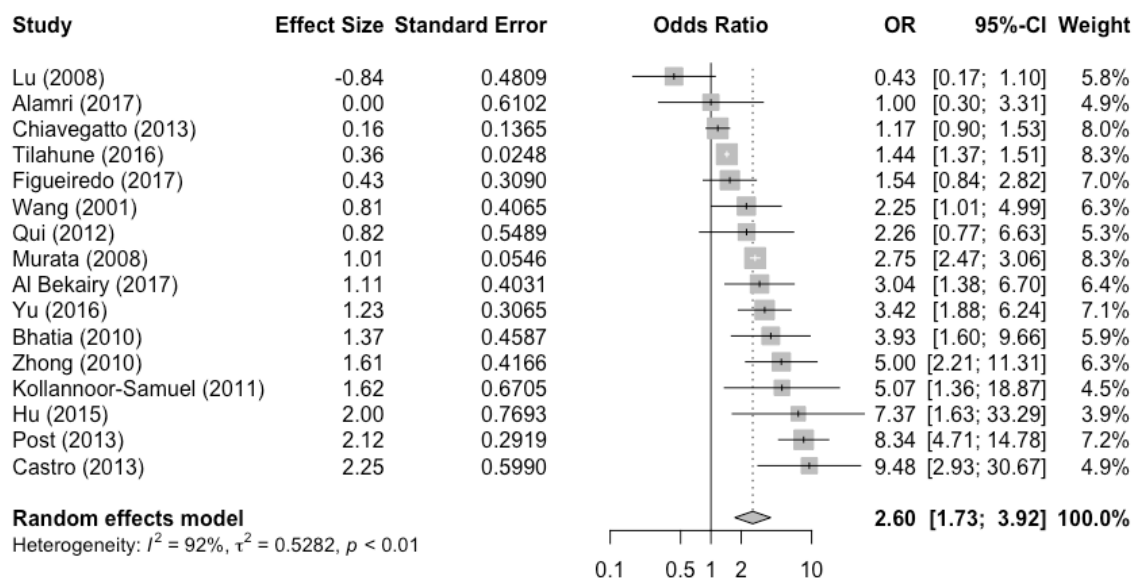


Figure 2.10: Forest plot of depression risk associated with low versus high income.

Results of the meta-analysis of depression risk in individuals with a medium income compared to individuals with a high income can be seen in Figure 2.11 below. The pooled estimate, calculated using a random effects model, show a 61% increase in risk of having depression for people with a medium income compared to people with a high income (OR = 1.61, CI = 1.24-2.09, $p = 0.0003$). A visual inspection of the individual studies in figure 2.9 reveals that five out of the 16 studies found an increased risk of depression for people with a medium income, 11 studies found a null result, and no studies found a reverse relationship between income and depression.

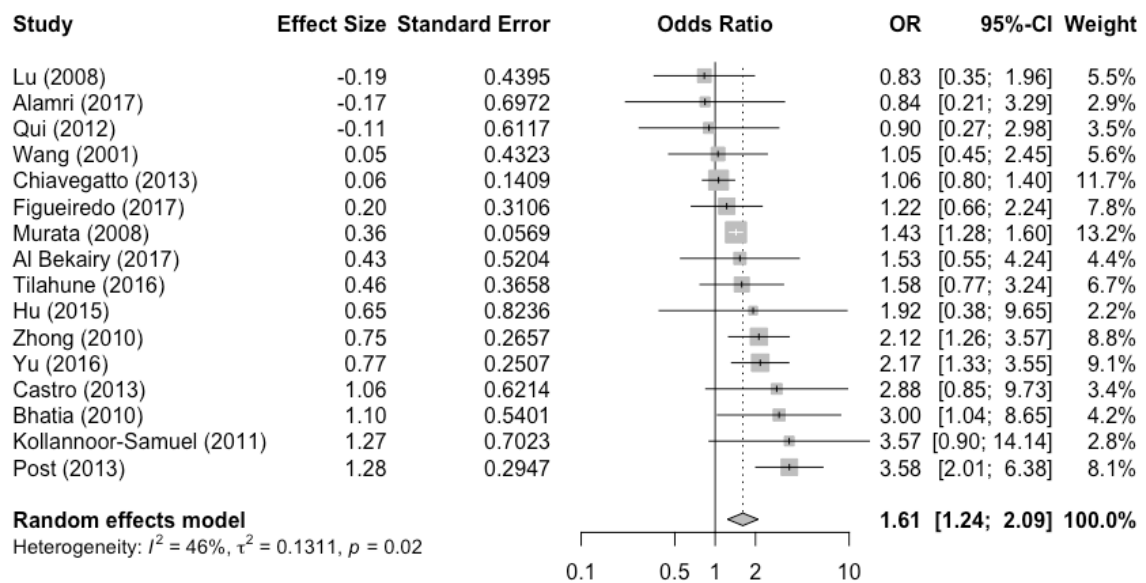


Figure 2.11: Forest plot of depression risk associated with medium versus high income.

The between study heterogeneity in the two meta-analyses of the association between low, medium, high income, and depression varies considerably. In the meta-analysis of low vs. high income groups the heterogeneity is estimated at $I^2 = 92\%$ with a variance of $\tau^2 = 0.52$. Whereas the meta-analysis of medium vs. high income revealed an I^2 of only 46% and a variance of $\tau^2 = 0.13$.

2.3.5.3 Educational attainment

The meta-analysis on depression risk across low, medium, and high educational attainment included 24 studies. A random effects model estimated the pooled risk of depression in individuals with a low (compared to a high) educational attainment to be at 81% (OR = 1.81, CI = 1.35-2.43, $p < 0.0001$). Results of the pooled estimate as well as risk estimates from each individual study can be seen in Figure 2.12. A visual inspection of the forest plot (fig. 2.10) shows that 11 studies found a significantly higher risk of depression in individuals with a low (compared to high) educational attainment, 12 studies found a null effect, and one study found the reverse relationship; a decreased risk of depression with low educational attainment.

Out of the 24 studies 2 ORs and CIs stand out, namely Kui (2014) and Hankerson (2015). Kui (2014) is the only study in this subgroup meta-analysis that

found a significantly increased risk of depression association with high educational attainment. Kui et al recruited participants from an epilepsy cohort with a high prevalence of depression (65 out of 215 participants) and only 10 out of 215 participants had a high educational attainment level. It is possible the clinical setting of recruitment, causing a high prevalence of depression, and a relatively small N have skewed the risk estimate. Hankerson (2015) recruited 122 participants from local churches, only seven of those participants had a low educational attainment whereas 51 had a high educational attainment. This skewed distribution may have resulted in the unusually high-risk estimate.

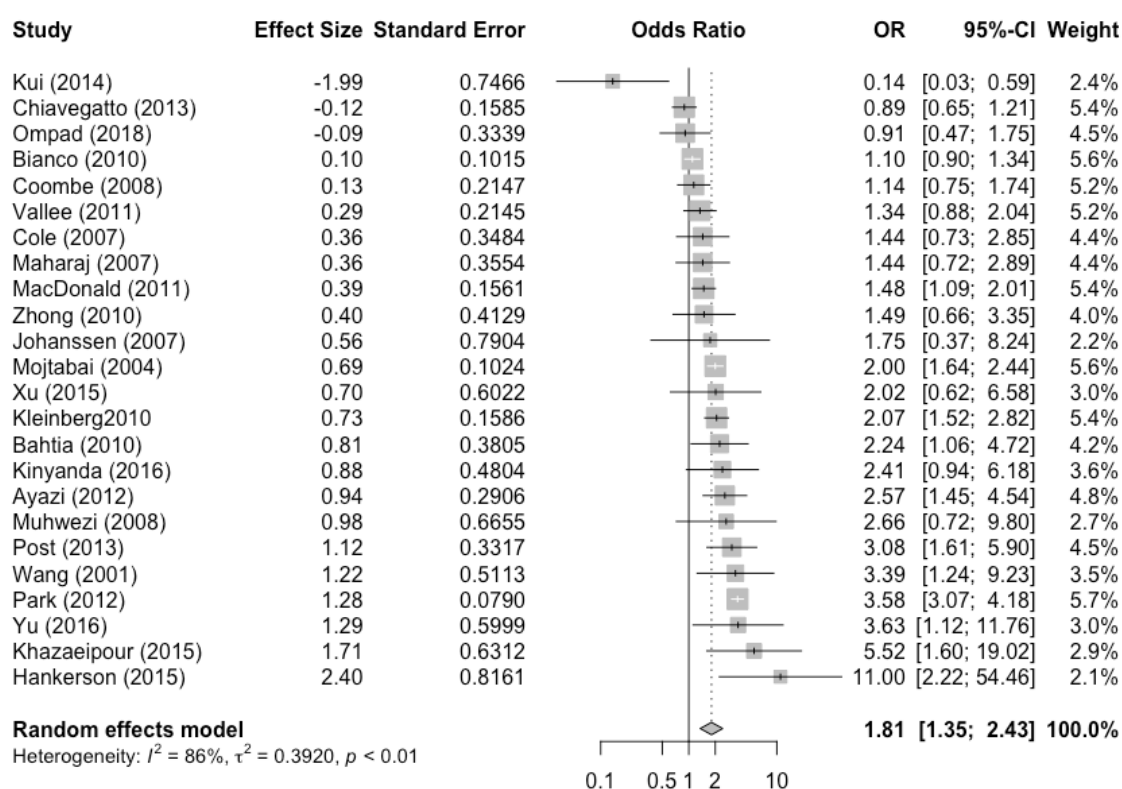


Figure 2.12: Forest plot of depression risk associated with low versus high educational attainment.

Results from the meta-analysis comparing risk of depression in individuals with a medium versus high educational attainment revealed a similar pattern. The random effects model revealed a pooled estimate of a 33% higher risk of depression in individuals with a medium educational attainment compared to people with a high educational attainment (OR = 1.33, CI = 1.06-1.67, $p = 0.013$). The results of the meta-analysis are visualised in a forest plot in Figure 2.13.

The risk estimates for the individual studies show that eight studies out of the 24 reported an increased risk of depression in individuals with a medium, compared to high, educational attainment, 16 of the studies found a null effect and none of the studies found a reverse relationship between medium versus high educational attainment and depression.

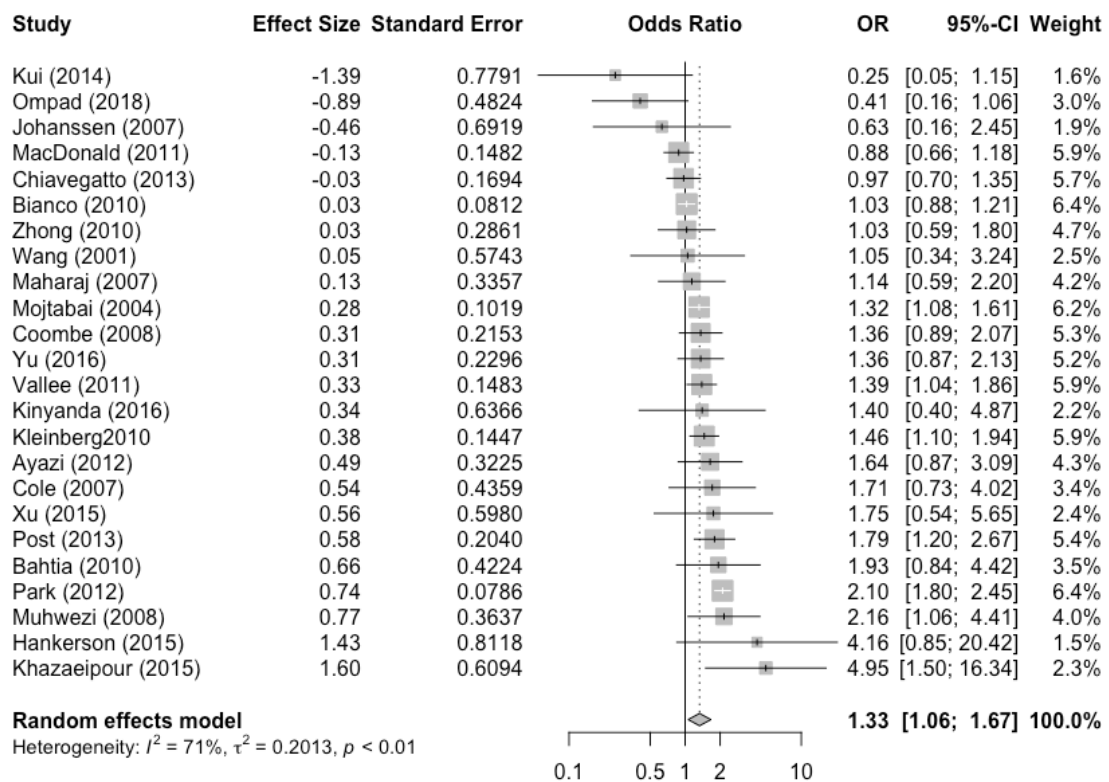


Figure 2.13: Forest plot of depression risk associated with medium versus high educational attainment.

Between study heterogeneity was relatively consistent across the two meta-analyses of low, medium, and high educational attainment, and depression. In the analysis of depression risk in low versus high educational attainment the heterogeneity was estimated at $I^2 = 86\%$ with a variance of $\tau^2 = 0.39$. The heterogeneity of the medium versus high educational attainment meta-analysis was estimated at $I^2 = 71\%$ with a variance of $\tau^2 = 0.20$.

2.3.6 Continuous data meta-analysis

Continuous results were reported by five included studies, three studies reported mean income and two studies reported mean number of years of education.

2.3.6.1 Income

The three studies reporting mean income were meta-analysed using Standardised Mean Differences (SMD) and a random effects model. The results of the meta-analysis can be seen in Figure 2.14, people with depression had a significantly lower mean income (SMD = -0.2855, CI = -0.4380 - -0.1330, p = 0.0002). Two studies found a lower mean income in the depressed group and one study found a null result. The between study heterogeneity is low, estimated at $I^2 = 48\%$ with a variance of $\tau^2 = 0.15$.

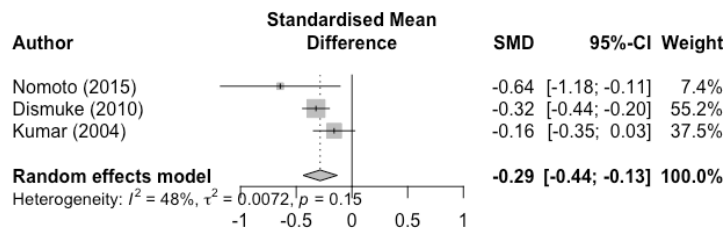


Figure 2.14: Forest plot of depression risk associated with income based on continuous data.

2.3.6.2 Education

The two studies reporting mean number of years of education were meta-analysed using Standardised Mean Differences (SMD) and a random effects model. The results of the meta-analysis can be seen in Figure 2.15, people with depression had a significantly lower mean number of years of education (SMD = -0.2869, CI = -0.5284 - -0.0454, p = 0.0199). Both studies found a lower mean number of years of education in the depressed group. The between study heterogeneity is low, estimated at $I^2 = 45\%$ with a variance of $\tau^2 = 0.18$.

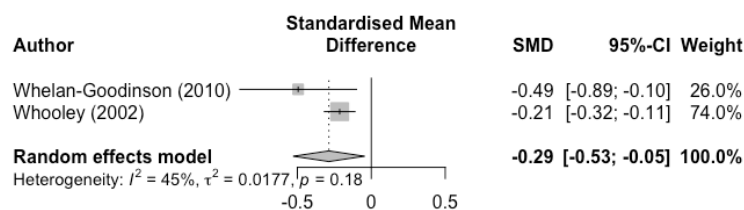


Figure 2.15: Forest plot of depression risk associated with educational attainment based on continuous data.

2.3.7 Longitudinal and narrative review

Fourteen longitudinal studies were included in the longitudinal review part of the systematic review. Of those studies, two reported results on SES, eight studies reported results on income, eight on educational attainment, and 3 on employment. Both longitudinal studies on SES and depression found a significant association between low SES and increased risk of depression. Low income was associated with increased risk of depression in six out of the eight longitudinal studies reporting results on income. A significant association between low educational attainment and depression was found in half of the eight studies. Finally, unemployment or low occupational group was associated with depression in all three studies, albeit in one of the studies it was only in women that the association was significant. A summary of the results of the longitudinal studies can be seen in Table 2.6.

Fifteen studies were included in the narrative review. Of those studies, 4 reported results on SES, 7 studies reported results on income, 10 on educational attainment, and 4 on occupation. Three out of 4 studies on SES found a significant association between low SES and increased risk of depression. Six out of 7 studies on income found a significant association between low income and increased risk of depression. Seven out of 10 studies on educational attainment found a significant association between low educational attainment and increased risk of depression. Half of the studies (2/4) on occupation found a significant association between low occupational status and increased risk of depression. None of the studies in the narrative review found a significant association in the opposite direction; high SES associated with increased risk of depression. Details and results of studies included in the narrative review can be seen in Table 2.7.

Author	Country	Recruitment setting	Years of follow-up	Sample size	N. of controls	N. of cases	Result
Lee (2016)	Taiwan	Community	10	1,743,948	1,727,693	16,255	Income: Percentage of low-income individuals with MDD (1.73%) significantly higher than non-low-income individuals (0.92%), $p < 0.0001$. No significant change in incidence rates across follow-up.
Stansfeld (2008)	United Kingdom	Community	3	9377	9238	139	Occupation: After adjusting for childhood SES women working in manual labour (compared to non-manual) were at increased risk of depression (OR = 2.53, CI = 1.59-4.04). No significant difference in risk of depression in men across occupation types.
Wang (2012)	Canada	Community	1	2752	2682	70	Low income (< \$60,000) was not significantly associated with increased risk of depression after follow-up (OR = 1.42, CI = 0.80-2.51). Low educational attainments (< university degree) was not significantly associated with increased risk of depression after follow-up (OR = 2.20, CI = 0.98-4.95).
Almeida (2012)	Australia	Community	2	18,758	18,483	275 (40 recurrent)	No post-school education was not significantly associated with baseline depression (OR = 1.1, CI = 0.7-1.7) or repeated depression at follow-up (OR = 0.7, CI = 0.3-1.8)
Joergensen (2016)	Denmark	Clinical	30 days (early) 2 years (late)	87,118	77,579	9,539	Low education (corrected for age and sex): Early depression HR (CI) = 1.13 (0.99-1.28). Late depression HR (CI) = 1.08 (1.03-1.13)

							Unemployment (corrected for age, sex, education): Early depression HR (CI) = 1.57 (1.29–1.90) Late depression HR (CI) = 1.53 (1.43–1.64) Income (corrected for age, sex, education, employment) was not significantly associated with early depression or late depression.
McGormick (2018)	USA	Clinical	1	848	682	166	Household income OR (95%) incident depression: <\$40,000: OR = 2.08 (1.40–3.10) \$40,000 to <\$80,000: OR = 0.92 (0.64–1.32) ≥\$80,000 (reference) Education, OR (95%) of incident depression: Did not graduate college: OR = 1.60 (1.13–2.26) College graduate (reference)
Anderson (2011)	Sweden	Community	5	30,970	28,500	2,470	Risk of depression associated with Socioeconomic status: Highest: ref. Second highest: 1.04 (0.86 - 1.26) Middle: 1.14 (0.95 - 1.36) Second lowest: 1.41 (1.17 - 1.70) Lowest: 1.53 (1.25 - 1.88)
Batterham (2009)	Australia	Community	4	6,605	6,334	271	Not in full time employment was associated with increased risk of depression 4 years later (OR = 2.1, X²= 8.3, p = .0040)

Hamano (2019)	Sweden	Clinical	7	188,907	166,893	22,014	Significant association between both educational attainment and income and development of depression in both men and women.
Kosidou (2011)	Sweden	Community	5	21,821	21,437	384	Significant association between income and depression in both men and women. No association between educational attainment and depression in either men or women.
Wirback (2018)	Sweden	Community	4	169,262	162,823	6,439	Disposable Household income High (ref.): Medium high: OR (95%) = 1.2 (1.1–1.2) Medium: OR = 1.3 (1.2–1.4) Medium low: OR = 1.2 (1.1–1.3) Low: OR (95%) = 1.0 (0.9–1.1) Parental education High (ref.): Medium: OR = 1.1 (1.1–1.2) Low: OR = 1.1 (1.0–1.2)
Zhao (2018)	China	Community	1	945	825	130	Significant association between poor economic situation and depression (OR = 8.294, CI = 3.12–22.02). No significant association between education and depression.
Tani (2016)	Japan	Community	50+	10,458	9,095	1,363	Childhood SES (at age 15) High: ref Middle: ARR = 1.13 (0.96–1.33) Low ARR = 1.44 (1.23–1.69)

Table 2.6: Longitudinal review of included longitudinal studies

Author	Country	Recruitment setting	Study design	N. of controls	N. of cases	Result
Turner et al. (1999)	USA	Community	Cross sectional	1239	154	Occupational level Major Professional: MDD % = 4.1 Lesser Professional: MDD % = 7.9 Minor Professional: MDD % = 9.7 Clerical/Sales: MDD % = 14.5 Skilled/Manual: MDD % = 7.2 Semiskilled/Unskilled: MDD % = 18.6
Wilson et al. (1999)	UK	Community	Cross sectional	1469	417	Townsend score: -1: OR (ref) = 1 0-1: OR (95% CI) = 2.64 (1.33-5.29) 4-5: OR (95% CI) = 1.30 (0.78-2.17) 6-7: OR (95% CI) = 1.28 (0.78-2.11) 8-9: OR (95% CI) = 1.89 (1.16-3.08) 10-13: OR (95% CI) = 2.0 (1.16-3.47)
Goodman et al. (2003)	USA	Community	Cross sectional	13602	1510	Parental income: Depression relative risk (95%CI) Q5: OR = 1 (ref.) Q4: OR = 1.21 (1.36, 1.95) Q3: OR = 1.46 (1.23, 1.74) Q2: OR = 1.63 (1.36, 1.95) Q1: OR = 2.07 (1.73, 2.47) Parental education OR (95%CI): Professional Degree: OR = 1 (ref.) College Graduate: OR = 1.5 (1.10, 1.88) > High School, < College: OR = 1.6 (1.31, 1.98) High School: OR = 1.95 (1.59, 2.40) < High School: OR = 2.95 (2.37, 3.67)
Beard et al. (2008)	USA	Community	Cross sectional		605	Household income \$100,000+: OR = 1.00 (ref.), \$75,000–\$99,999: OR (CI) = 1.27 (0.69–2.33), \$50,000–\$74,999: OR (CI) = 1.67 (0.96–2.88), \$40,000–\$49,999: OR (CI) = 3.46 (1.92–6.22), \$30,000–\$39,999: OR (CI) = 4.22 (2.39–7.45), \$20,000–\$29,999: OR (CI) = 3.39 (1.96–5.86),

						<\$20,000: OR (CI) = 5.05 (3.04–8.41)
						Educational attainment Graduate work: OR = 1.00 (ref.), Bachelor of arts degree: OR (CI) = 1.46 (0.90–2.39), Some college: OR (CI) = 1.99 (1.20–3.32), High school/GED: OR (CI) = 1.97 (1.19–3.29), <High school: OR (CI) = 3.22 (1.83–5.69)
Pikhartova et al. (2009)	Czechia	Community	Cross sectional	6123	1493	Deprivation level, OR (95% CI): 0: OR = 1 (ref) 1–2: OR = 1.91 (1.61–2.26) 3–5: OR = 2.76 (2.36–3.23) 6–12: OR = 4.87 (4.07–5.82)
						Education, OR (95% CI): Primary or less: OR = 1 (ref) Vocational: OR = 0.73 (0.62–0.86) Secondary: OR = 0.63 (0.53–0.75) University: OR = 0.48 (0.38–0.62)
Poongothai et al. (2009)	India	Community	Cross sectional	21608	3847	Income level, OR (95% CI): > 20000: OR = 1 (ref.) 10001–20000: OR = 1.15 (0.54–2.43) 5000–10000: OR = 1.41 (0.70–2.81) < 5000: OR = 2.44 (1.24–4.81)
						Education status, OR (95% CI): Professional: OR = 1 (ref) Post graduate: OR = 1.09 (0.52–2.31) Graduate: OR = 1.35 (0.70–2.59) SSC: OR = 2.00 (1.05–3.82) Below SSC: OR = 2.66 (1.40–5.07) Illiterate: OR = 4.77 (2.50–9.11)
Braich et al. (2011)	India	Community				Income: <2500: OR = 1 (ref.) 2500–6000: OR = 0.31 6000–15000: OR = 0.17 15000–24000: OR = 0.06 >24000: OR = 0.68

Andersen et al. (2011)	Brazil	Community	Longitudinal (6yrs)		T1: 2% T2: 4.9%	Percent depressed per employment group: 2000 Non-manual = 0.7% 2000 Manual = 1.3% 2000 Non-employed = 11.3% 2006 Non-manual = 2.8% 2006 Manual = 4.7% 2006 Non-employed = 16.8%
Yunming et al. (2012)	China	Community	Cross sectional	1085	401	Max income (RMB) 0–300: OR = 1(ref) 301–500 : OR = 1.11 (0.21–5.79) 501–1000 : OR = 1.12 (0.26–4.90) 1001–1500: OR = 0.73 (0.20–2.66) 1500 or above: OR = 0.42 (0.15–1.16) Education level No education: OR = 1 (ref) Elementary school: OR = 0.96 (0.13–7.27) Primary school: OR = 1.34 (0.31–5.71) Secondary school: OR = 1.40 (0.27–7.23) Post-secondary school/above: OR = 0.85 (0.28–2.62)
Alonso-Moran et al. (2014)		Clinical	Cross sectional		12392	Deprivation index (95% CI) 1: OR = 1 2: OR = 0.95 (0.88, 1.03) 3: OR = 1.01 (0.93, 1.09) 4: OR = 0.93 (0.86, 1.01) 5: OR = 0.96 (0.89, 1.03)
Shi et al. (2015)	China	Clinical	Cross sectional	3639	3800	Social status (%) Semi skilled and unskilled workers: MDD = 9.21, Control = 11.31 Skilled manual employees: MDD = 28.52, Control = 26.2 Administrative personnel, minor professionals, clerical and sales workers: MDD = 17.24, Control = 14.1 Executives, business owners, major professionals: MDD = 27.45, Control = 29.07 Other: MDD = 17.57, control = 19.31 Education (%): no education or pre-school: MDD = 6.82, Control =

						3.97 primary school or below: MDD = 18.05, Control = 12.95 junior middle school: MDD = 31.27, Control = 32.92 senior middle school : MDD = 29.24, Control = 30.26 junior college: MDD = 8.85, Control = 11 bachelor degree: MDD = 4.92, Control = 7.89 master degree or above: MDD = 0.85, Control = 1
Magklara et al. (2015)	Greece	Community	Cross sectional	2181	246	Father's employment status OR (95%): Public sector employee: 1.00 (ref.) Private sector employee: 1.29 (0.85-1.95) Self-employed: 0.92 (0.62-1.37) Retired: 0.41 (0.17-1.00) Unemployed/ Other: 1.02 (0.58-1.80) Father's educational level OR (95% CI): Primary: 1.00 (ref) Secondary Basic: 1.03 (0.56-1.85) Secondary Complete: 1.01 (0.63-1.62) Technological degree: 1.41 (0.81-2.45) University degree: 0.94 (0.57-1.54)
Freeman et al. (2006)	Spain Finland Poland	Community	Cross sectional	Finland = 1854 Poland = 3726 Spain = 4073	Finland = 80 Poland = 214 Spain = 510	SES (0–110) index Finland: OR (95% CI) = 0.991 (0.98–0.997) SES (0–110) index Poland: OR (95% CI) = 0.986 (0.975–0.99) SES (0–110) index Spain: OR (95% CI) = 0.989 (0.984–0.995) Income Quintile Finland: OR (95% CI) = 0.84 (0.714–0.987) Income Quintile Poland: OR (95% CI) = 0.84 (0.71–0.98) Income Quintile Spain: OR (95% CI) = 1.0 (0.914–1.09) Yrs of Education Finland: OR (95% CI) = 0.94 (0.89–0.985) Yrs of Education Poland: OR (95% CI) = 0.934 (0.88–0.983)

						Yrs of Education Spain: OR (95% CI) = 0.913 (0.887–0.939)
Altino et al. (2017)	Brazil	Clinical	Cross sectional	97	23	Mean number of years of education: Depressed: 6 (4–8) Not-depressed: 11 (4–15) p = 0.011
Lotfaliany et al. (2018)	China Ghana India Mexico Russia South Africa	Community	Cross sectional	39302	2508	Wealth quantile, OR (95% CI) Q1: OR = 1.57 (1.33,1.84) Q2 : OR = 1.50 (1.28,1.76) Q3 : OR = 1.43 (1.22,1.67) Q4 : OR = 1.16 (0.99,1.36) Q5: OR = 1 (Ref.) Education, OR (95% CI) No formal education: OR = 1.71 (1.31,2.22) Less than primary level: OR = 1.68 (1.28,2.21) Primary level: OR = 1.33 (1.01,1.74) Secondary level: OR = 1.33 (1.04,1.71) Tertiary level: OR = 1 (Ref.)

Table 2.7: Details and results of studies included in the narrative review.

2.4 Discussion

The aim of the present systematic review and meta-analysis was to investigate a possible association between depression and socioeconomic status (SES), to provide an estimate on a possible difference in risk across different levels of SES, and to estimate this estimate in the largest possible empirical evidence base of well-conducted studies.

This systematic review on the association between depression and SES is, with 138 studies included, to date the largest of its kind. The studies span almost five decades of research, 47 countries across six continents, covering the whole human life span, and includes 34 different comorbid diseases.

2.4.1 Main findings

In all but one of the nine separate meta-analyses a consistent association between low SES and depression was found. SES measures, and each subpart of SES (income, education, and employment) were meta-analysed separately.

In the meta-analyses of SES and depression a significant association between low SES and increased risk of depression was found in two out of the three analyses. A pooled estimate from studies reporting dichotomous results ($N = 15$) revealed a 70% increased risk of depression in people with a low SES compared to people of high SES. A pooled estimate from studies reporting ordinal results ($N = 7$) showed a 120% increased risk of depression in people with a low SES. However, a meta-analysis of depression risk in people with medium, compared to high, SES showed no significantly increased risk ($OR = 1.40$, $CI = 0.84-2.33$, $p = 0.195$).

When meta-analysing EA, a significantly increased risk of depression in people with low compared to high EA (or medium EA compared to high EA) was found across all four meta-analyses (dichotomised-, ordinal-, and continuous results). Out of the 80 studies included in the meta-analyses of EA and depression two reported a reverse relationship, associating high EA with an increased risk of depression (Kim et al 2014, Hui et al 2014).

Results on the association between income and depression were reported, and meta-analysed with dichotomised, ordinal, and continuous data. Across all four meta-analyses a significantly increased risk of depression was found in people with a low or medium income,

compared to a high income. None of the 44 studies included in the meta-analyses on income reported a significant reverse relationship between income and depression.

Unemployment was significantly associated with a 57% increased risk of depression in the meta-analysis of 29 studies reporting dichotomous results.

Eligible longitudinal studies were included in the narrative part of the systematic review. The 14 studies reported 21 investigations into the association between SES and depression and of those, 15 found a significant relationship between low SES and depression whilst seven comparisons revealed a null result. None of the investigations found a significant relationship between high SES and depression. The most inconsistent SES and depression association was to be found in the educational attainment subgroup where half the comparisons revealed a significant association between low EA and increased risk of depression while the other half of the studies revealed a null finding.

Overall, the findings from the longitudinal studies and the cross-sectional studies are relatively congruent; the majority of studies report a significant association between low SES and increased risk of depression.

The majority of studies found a significant association between low SES and increased risk of depression, and a few studies found a null result, just three studies found a reverse relationship, i.e. low SES being associated with a decreased risk of depression.

Kim et al. (2014) who found that people with less than high school (compared to more than high school) education were less likely to have depression (OR = 0.31, CI = 0.23-0.41, N = 2,355). In the same study unemployment seemed associated with higher risk of depression but this was not significant (OR = 1.26, CI = 0.98-1.62, N = 2,119). Overall, the study has a high prevalence of depression (36.1% with CES-D). Participants were older adults (65+) recruited from a rural county-based health promotion program. The high prevalence is not unusual for an older South Korean population as was reported by Park et al (2012) included in this review. Furthermore, Kim et al. point out that in South Korea there is a tendency to report somatic symptoms as emotional disturbances which may have influenced the CES-D scoring. Finally, the authors speculate that people with a higher EA may live in more competitive environments which in turn causes psychological distress. It is worth pointing out that Park et al. (2012), a study from South Korea with 4,355 participants with a mean age of 73, did not find a similar result to Kim et al. (2014). They did indeed find that low educational attainment was associated with increased risk of depression.

Kui et al. (2014) also found a reverse relationship between EA and depression, more specifically low EA (compared to high EA) was associated with a decreased risk of depression (OR = 0.14, 0.03-0.59). The low EA and decreased depression risk association was not significant when comparing risk of depression in people with medium EA to high EA. The study involved 215 Chinese epilepsy patients, 65 of which were diagnosed with major depression. Educational attainment was divided into three groups (low, medium, and high) and a chi squared test did not reveal a significant difference in EA across the two groups. The aberrant, for this meta-analysis, direction of depression risk observed in this study may therefore be due to the step-wise meta-analysis of risk in studies reporting three SES groups. Finally, the effect could be due to the relatively large proportion of an otherwise small overall number of university graduates with depression, in this study (7 out of 10).

Finally, Bekaroglu et al. (1991) found a significantly lower risk of depression in people with medium SES compared to high SES (OR = 0.39, 0.16-0.96) however there was no effect of SES on depression risk when comparing low and high SES. Bekaroglu et al. included 292 individuals from Turkey, aged over 60. The authors do not comment on the association between SES and depression, but it is worth noting that the number of people with high SES (N = 23) is a lot smaller than then number of people with low SES (N = 114) and medium SES (N= 155). Once again, this aberrant direction of depression risk may be a result of a comparatively smaller high-SES group and the step-wise meta-analysis of depression risk between the three SES categories employed in this review.

The meta-analyses of dichotomised data revealed an increased risk of depression associated with each one of the low SES measures (SES, income, education, income, and employment). The risk estimates are relatively uniform across the SES measures with ORs ranging from 1.57 (employment) to 1.95 (income). In particular, employment, education and SES have similar risk estimates, and the 95% CIs of employment (OR = 1.57, 1.39-1.81) and education (OR = 1.69, 1.43-2.00) are fully encompassed by that of SES (OR = 1.70, 1.32-2.19). From a theoretical perspective this makes sense as SES is a composite measure made up of components either directly or indirectly related to income, education, and employment. Furthermore, the three subcategories of SES will inevitably share a substantial level of variance. Income, education, and employment are, to varying degrees, intercorrelated whilst also each accounting for a proportion of unique variance. Having a high level of education makes it more likely to obtain a high income, whilst having a high income makes it easier to

retrain or pursue a higher education. Having any income at all often relies on being employed, although this does differ from country to country depending on the strength of the social security. However, these relationships are far from set in stone and many societal and individual factors influence the strength of the relationships between the SES components. When partitioning the risk of depression associated with SES into component risk factors it is important to remember that these components are correlated with one another.

Income shows a stronger association with depression risk than the rest of the SES components. With a 95% increased risk of depression in people with a low-, compared to high income (OR = 1.95, 1.51-2.52). The meta-analyses of studies dividing SES measures into three groups further support this finding. The meta-analyses of depression risk at low versus high income and medium versus high income revealed the highest ORs of any of the ordinal meta-analyses. Our sensitivity analysis showed no significant difference in the risk estimates for high income economies and low-income economies, suggesting the strong association between income and risk of depression is relatively consistent across cultures. The reason for this stronger observed effect is unclear from the present data. It is also important to point out that the abovementioned estimates do not give an indication of directionality. From the present meta-analysis it is not known whether low income increases the risk of depression or whether depression increases the risk of having a low income. However, six out of eight longitudinal studies showed an increased risk of depression following low income, the other two found no significant association, suggesting the direction of causality may be from low income to increased risk of depression.

The reason behind the difference in risk estimates of the SES measures, is unclear. It is possible that low income, or poverty, is particularly restrictive in terms of direct and indirect health promoting behaviours as well as increasing psychological stress. It is also possible that income has a more direct effect on depression risk as it is often the knock-on effect of other SES measures. Low education and/or unemployment often lead to relatively lower income and therefore income may be closer to depression in the chain of effects. This finding could be particularly important for informing policy intervention attempts. Of all the SES measures, income is the one that is the easiest to change with an almost immediate effect. Changing someone's overall socioeconomic status takes a long time and is difficult, getting re-educated can take many years and changing employment rates often involves changing the structure of the labour market through macro-economic measures or getting the

individual ready for the existing labour market which entails removing the individual barriers in place such as health, carer commitments or re-training.

An increasingly appealing intervention against poverty is universal basic income (UBI) ([Arnold 2018](#)). At its core UBI means giving a regular amount of money to an individual without any obligations or commitments necessary. Smaller trials of this intervention have had positive effects on health, wellbeing and SES (Enns et al., 2021; Forget et al., 2011; Koistinen & Perkiö, 2014). Many of these interventions are relatively short-lived so it is difficult to predict the impact they might have in the long run. However, larger, and longer interventions are underway in Kenya and the United States lasting up to 12 years. Studies on the SES and depression association are often of an observational nature and as such are prone to biases. However, using these observational studies to inform the design of interventional experimental studies could be a methodologically strong approach. Given the present systematic review found the highest risk of depression to be associated with low income, UBI may present a great preventative opportunity in an intervention based experimental set-up.

2.4.2 Strengths and limitations

The main limitation of the present systematic review is the high levels of heterogeneity in the meta-analyses, which may in-turn reflect differences in study populations, methods or other factors (e.g. cohort effects where the association effect size changes over decades). The analyses involving dichotomous data all have heterogeneity estimates (I^2) between 81% and 89%. Sensitivity analyses were performed on each of the meta-analyses to investigate if differences in countries' economies could explain the high-level of heterogeneity. None of the meta-analyses differed significantly in risk estimates between low-income economies and high-income economies. The sensitivity analyses of income and unemployment results revealed that most of the heterogeneity were in studies from high-income economies. However, for both meta-analyses there were four times as many studies in the high-income economy group which may be behind the difference in between-study heterogeneity.

A good example of why the heterogeneity may not stem from differences in cultures can be found in one of the studies included in this review; Ambugo et al (2014). The study compares SES and depression associations in four countries (Ghana, Kenya, UK, and

Norway) spanning low- and high-income economies, but crucially all with a similar design. The risk estimates are similar and a meta-analysis of the effects reveals relatively little heterogeneity ($I^2 = 50.9$, $Q = 6.11$).

The high levels of between study heterogeneity could instead be due to variation in depression diagnosis screening tools across studies, a total of 24 different diagnostic instruments were used in this systematic review, with the Centre for Epidemiological Studies Depression (CES-D) being the most prevalent screening tool. The setting of recruitment could also influence the level of between study heterogeneity. Studies recruiting in a clinical setting often have a much higher prevalence of depression compared to community recruited studies. Finally, the difference in covariates across studies could account for some of the heterogeneity. Most studies did not include covariates, but some studies controlled for, for instance, age and sex in their analyses.

Another limitation of the current study is the lack of longitudinal studies in the meta-analysis. This was done because of the low number of included longitudinal studies and the large variations in follow-up time. However, the most recent review on SES and depression focused solely on longitudinal studies and the narrative review of longitudinal studies revealed similar results to those of the meta-analyses.

In systematic reviews of observational data it is difficult to confidently infer causality even when including longitudinal data. In the case of investigating the association between a complex disease and SES this is even more pronounced. As mentioned in the general introduction there are myriads of factors influencing an individual's SES (both environmental and genetic) and as a result ascribing causality to a single factor in an SES- depression association is difficult. Nevertheless looking at the relative differences in risk contribution between SES factors may give an insight into the dynamics underlying to association between SES and disease risk.

A final limitation of the current systematic-review is the lack of a formal assessment of the quality of studies. As mentioned earlier in this chapter, it is difficult to formally assess and quantify the quality of studies when dealing with observational/prevalence designs and attempting to use less than ideal metrics could at best create a false sense of confidence and at worst introduce bias.

The strengths of the present systematic review are its cross cultural, cross age-group, and cross morbidity estimates of the association between SES and depression, as well as providing risk estimates modelled on the largest evidence base to date.

Furthermore, the present meta-analyses are unlikely to be hindered by publication bias, and as a consequence less likely to over-estimate the real risk of depression associated with SES. For many of the included studies, the SES and depression association was not their primary outcome, this may also explain why a number of studies have reported null findings.

Finally, the inclusion of all measures of SES, whether composite or single indicator measures, gives an opportunity to compare the risk of depression across subparts of SES. This, in turn, lays the groundwork for future research to pinpoint which part of an individual's SES confers the highest risk of depression.

2.4.3 Future work

Future research should strive to elucidate the direction of causality in the SES-depression association. With limited resources and for interventions to have an effect. We need to know which individuals to focus on and where to invest. If depression causes low SES it is important to support people with a diagnosis of MDD to maintain or even improve their SES. On the other hand, if low SES causes depression, it is important to support people with low SES and focus on early detection and intervention of depression in this demographic.

One way of doing this is to setup cross generation longitudinal population cohort studies, an existing candidate for such a study could be the Avon Longitudinal Study of Parent and Children (ALSPAC) with data on several generations of participants.

Another tool for examining causal directions is Mendelian Randomisation (MR) models. Previous MR studies into the association between low SES and depression found a significant association between the two factors but did not find a robust indication of causal direction (Harrison et al., 2020). Future MR studies would benefit from stronger genetic instrumental variables and should investigate the possibility of a bidirectional relationship between low SES and depression.

A better understanding of genetic risk of depression and the genetic make-up of SES, will also provide the opportunity to tease apart what is the social or environmental part of SES that confers risk of depression and which part is genetic SES risk. This information could aid in pinpointing areas of maximum impact for interventions or even help guide research into preventative treatments.

Hopefully the present systematic review and resulting risk estimates will help to convey the importance of including social measures, such as SES, in psychiatric research. To understand how depression develops we need to include SES in our models, not only to remove residual confounding but to understand how it is contributing to overall risk of depression.

3 DNA Methylation Signatures of Area-Based Deprivation

3.1 Introduction

Socioeconomic status (SES) traditionally refers to an individual's level of income, education and their occupational status and type. It is often used as a proxy for overall material deprivation or social class, and may be measured at the individual level or estimated from neighbourhood-wide figures.

Low SES has repeatedly been associated with decreased life expectancy in general (Lleras-Muney, 2005), (Stringhini et al., 2017) and more specifically with an increased risk of many physical diseases and mental health conditions (Everson et al., 2002). In the UK, life expectancy in women in the 10% most deprived areas has decreased over the past decade and people in more deprived areas spend a larger proportion of their already shortened lives in ill-health (Marmot report 2020). Low SES has been associated with higher rates of cardiovascular disease (Diez-Roux et al., 1997) chronic obstructive pulmonary disorder (COPD) (De Matteis et al., 2016; Prescott & Vestbo, 1999), asthma (Kozyrskyj, Kendall, Jacoby, Sly, & Zubrick, 2010), and diabetes (Marmot et al., 1991) A similar picture has emerged for mental health problems including depression (Bruce, Takeuchi, & Leaf, 1991), schizophrenia (Agerbo et al., 2015; Werner, Malaspina, & Rabinowitz, 2007) and anxiety (M. J. Green & Benzeval, 2013).

The association between low SES and ill health, both physical and mental, can be partly attributed to lifestyle factors, in particular; smoking, increased levels of obesity (BMI: Body Mass Index), and increased alcohol consumption (Lynch, Kaplan, & Salonen, 1997). The significant association of low SES and all-cause mortality is, however only partly explained by these lifestyle factors. After covarying for smoking, alcohol consumption, and physical inactivity a mortality risk ratio of 1.59 remains for low SES (Nandi et al., 2014). These results highlight the clinical and scientific need to better understand the association between low SES, ill-health, and the underlying biology, over and above the effects of known life-style risk factors.

Recent genome wide association (GWAS) studies (Okbay et al., 2022) have identified over 3,000 single nucleotide polymorphisms (SNPs) associated with educational attainment and have shown that combining disease specific polygenic scores (PGS) and educational attainment PGS increases the predictive power for several diseases, among others cardiovascular diseases, diabetes, rheumatoid arthritis, asthma, and major depression. However, these findings give an indication of the innate biological risk of low SES rather than indications of the biological impact of living in low SES and its effect on health.

An alternative approach to understanding the relationship between SES and health is to examine DNA methylation (DNAm). DNAm is a type of epigenetic modification which regulates the transcription and splicing of genes. During methylation a methyl group is added to the DNA strand typically at cytosine-phosphate-guanine (CpG) sites. Importantly levels of DNAm changes with environmental exposures in a non-random manner and has been documented to change in a dose dependent relationship (McCartney et al., 2018). DNAm has been shown to be associated with many lifestyle factors associated with lower SES (McCartney et al. 2018, Lohoff et al. 2021, Chen et al. 2021), as well as with its potential consequences (Ligthart et al., 2016; Marttila et al., 2015; McDermott et al., 2016).

Previous research has found suggestive links between low SES and differences in DNAm measured in leukocytes in a selected subset of CpGs ($n = 110,631$) in adults (McDade et al., 2019; Stringhini et al., 2017) and in neonates (Coker, Gunier, Huen, Holland, & Eskenazi, 2018) as well as global DNAm levels in leukocytes (McGuinness et al., 2012). However, to date there are no whole blood, epigenome-wide explorations of differences in DNAm associated with SES. The benefit of an epigenome wide study of low SES is the potential for an unbiased account of the impact of low SES on the molecular level across the entire epigenome. Furthermore, it can show both detailed CpG-level differences in methylation as well as information on differences in methylation at the biological pathway level.

Low SES is associated with an increased risk of numerous diseases and disorders, among them depression. The last chapter found a robust association between low SES and increased risk of depression across SES subdomains and across cultures. However, the biological impact of living with low SES, over and above the negative effects of obesity, alcohol consumption and smoking, is poorly understood. To identify novel biological pathways underlying the relationship between low SES status and general health, in particular

mental health, an epigenome wide association study of area-based multiple deprivation was performed. The analysis makes use of deep phenotypic data to covary for three known environmental risks; alcohol consumption, smoking behaviour, and BMI. The study leverages the single largest DNAm dataset; Generation Scotland: Scottish Family Health Study (GS:SFHS), including more than 7,000 participants across the whole deprivation spectrum.

3.2 Methods

Phenotypic and DNA methylation data were obtained from the family-based population cohort study Generation Scotland (between 2006 and 2011) from across Scotland with both cross sectional and longitudinal data (B. H. Smith et al., 2013). Generation Scotland, set up as a collaborative cohort between Scottish medical schools and the National Health Service, originally recruited participants between the ages of 35 and 65, this was later widened to include participants between the ages of 18 and 96. Potential participants were identified at random from general medical practices across Scotland. Each contacted participant was asked to identify a family member that would be eligible to take part as well. A total of 126,000 people were invited to join the cohort, of which 12.3% agreed to participate and met inclusion criteria. In total 6,665 people agreed to take part and 16,007 of their family members similarly agreed to participate in the Generation Scotland cohort. Slightly less than half the participants (39%) lived in areas above average level of multiple deprivation, the prevalence of depression (4% vs. 8%), chronic pain (32% vs. 46%), heart disease (5.6% vs. 7.3%), and diabetes (4.1% vs. 5.7%) were all lower than the Scottish population prevalence. The present study included 7,702 individuals for whom data was available on DNAm, area-based deprivation, smoking behaviour, BMI, alcohol consumption and other covariates, see below. Baseline data was collected between 2006 and 2011.

GS:SFHS received ethical approval from NHS Tayside Research Ethics Committee (REC reference number 05/S1401/89) and has Research Tissue Bank Status (reference: 15/ES/0040). Written informed consent was obtained from all participants.

3.2.1 Phenotypes

An area-based measure of deprivation, the Scottish Index of Multiple Deprivation (SIMD), was used in the present study. The SIMD is a composite measure made up of seven

domains: income, employment, health, education, housing, geographic access, and crime. Scotland is then divided into geographical areas called data zones of roughly equal population; each data-zone includes between 700 and 800 people. That equates to 6,976 data zones across Scotland, all of which are ranked from 1, being the most deprived, to 6976 being the least deprived. This means it is a relative measure and the distance between each rank is not constant.

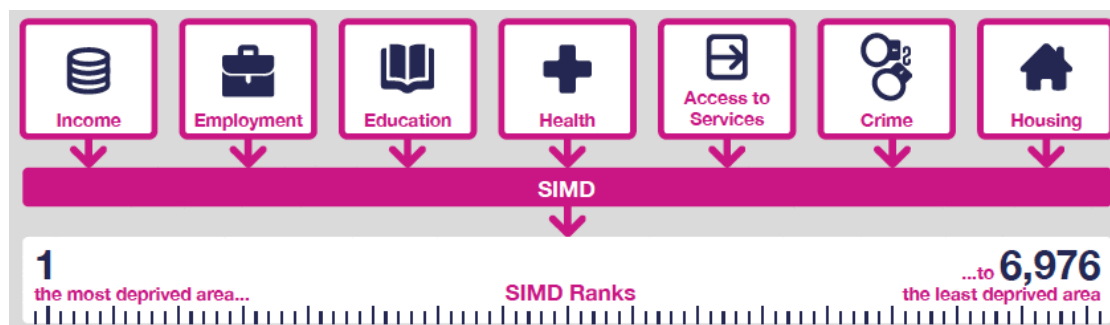


Figure 3.1: Illustration of the seven domains contributing to an area's SIMD rank, as well as the direction of ranking (<https://www.gov.scot/publications/scottish-index-multiple-deprivation-2016/>)

SIMD rank was used as a SES measure because it includes more information on the environment than other SES measures. Other potential measures include a composite SES score, subparts of SES such as income, educational attainment, or employment status. Townsend deprivation index, incorporating non-car ownership, non-homeownership, unemployment, and household overcrowding (Townsend, 1987), could also have been used as an SES measure. However, none of those SES measures capture as much information about the physical lived environment as SIMD. It was deemed important to capture as much information as possible about the level of deprivation in an individual's lived environment, when trying to investigate the consequences it has on an individual's biology.

Each domain is weighted individually the employment and the income domain are weighted the heaviest, followed by the education and health domain, then the access domain, then the crime domain and finally the housing domain. The Employment domain includes information on unemployment/disability benefit claimants in the data zone. The income domain counts the number of adults and children in receipt of tax credits as well as households claiming income support. The education domain incorporates information on school pupil attendance and performance, the proportion of the working age population

without a qualification, young people in higher education versus not in education employment or training. The health domain includes information on mortality ratio, hospital stays related to alcohol or drug misuse, emergency stays in hospital, prescriptions of anxiety, psychosis or depression medication, and the occurrence of low birthweight. The access domain is divided into two; the drivetime subdomain and the public transport subdomain each of which includes a calculation of time taken to a GP, retail centre, a school, and the post office. The crime domain includes numbers of house break-ins, drug offences, common assaults, violence, vandalism, and sexual offences. The housing domain counts the number of people in over-crowded households and the number of people in households without central heating (Shaw et al., 2007).

Body mass index (BMI) was calculated using height (cm) and weight (kg) measurements recorded by clinical staff at baseline visit. Alcohol consumption is represented by the number of units consumed in the week prior to the baseline visit. We used two measures of smoking, namely smoking status and pack years. All three measures were recorded with a self-report questionnaire at the baseline visit. For smoking status participants indicated whether they currently smoke, used to smoke but stopped within the last 12 months, used to smoke but quit more than 12 months ago or never smoked. Pack years were calculated by multiplying the number of cigarettes smoked per day with the number of years the participant has smoked. A similar approach has previously been successfully adopted by studies using DNAm data from Generation Scotland (Miruna C Barbu et al., 2021; Madden et al., 2021).

3.2.2 DNA methylation

Epigenome-wide DNAm was obtained from using Illumina HumanMethylationEPIC (850k) BeadChip array. Data collection was performed in two waves, while pre-processing and quality control was done in one batch. The entire Generation Scotland DNAm dataset (both data collection waves) was used for the EWAS while the two data collection waves were used separately as a training and testing dataset respectively for a deprivation-prediction replication analysis.

Standard processing procedures were used, similar to those employed by previous studies using Generations Scotland DNAm data (Barbu et al 2021). Pre-processing was done

in R version 3.5.1. The minfi package (Aryee et al., 2014) was used to read the raw datafiles. The shinyMethyl (Fortin, Fertig, & Hansen, 2014) package was used to exclude outliers, to plot the log median intensity of methylated vs. unmethylated signal for each array, and to exclude samples where methylation-predicted sex differed from recorded sex. waterMelon (Schalkwyk, Pidsley, Wong, & Touleimeimat, 2012) was used to remove samples where 1% or more of CpGs had a detection p-value of less than 0.05, probes with a bead count less than 3 in more than 5% of the samples and to remove probes for which 0.5% or more of samples had a detection p-value of less than 0.05. waterMelon was also used to normalise the data. Subsequently lumi (Du, Kibbe, & Lin, 2008) was used to convert the data into M-values. M values are calculated as the log₂ ratio of the intensity of methylated probes versus unmethylated probes. M- values are used to satisfy the assumption that the data is homoscedastic. Beta- values are interpreted more intuitively but the standard deviation of Beta-values is not static, i.e. Beta-values introduce heteroscedasticity As a consequence M-values are used in the methylation analysis and Beta-values will be used when reporting and comparing results (Du et al., 2010).

DISSECT (Canela-Xandri, Law, Gray, Woolliams, & Tenesa, 2015) was used to pre-correct these M-values for estimated blood cell types (granulocytes, natural killer cells, B-cells, CD4⁺ T-cells and CD8⁺ T-cells), relatedness and batch number.

Suboptimally binding or cross-hybridising probes as predicted by (McCartney et al., 2016) or (Zhou, Laird, & Shen, 2017) as well as X or Y chromosome probes were removed. Data from participants whose DNAm data indicated an XXY genotype and who had answered “yes” to all self-reported conditions were excluded.

The final dataset comprises corrected M-values at 713,522 loci for 7,702 participants. Data collection and pre-processing of the DNAm data was performed by the another team at The University of Edinburgh. A large proportion of that pre-processing was done by Dr Rosie Walker.

3.2.3 Statistical methods

The open access statistical software programme R version 3.5.1 was used for all statistical analyses and data visualisations.

Epigenome wide associations were carried out by computing a linear regression for each CpG probe on a chromosome-by-chromosome basis. Each model regressed DNAm

levels over SIMD rank correcting for smoking (pack years and status), BMI, alcohol consumption, age, sex, white blood cell estimates and 20 principal components to correct for unmeasured confounders:

$$CpG \sim SIMD\ rank + smoking\ status + pack\ years + BMI + alcohol\ consumption + age + sex + 20\ PCs$$

Any CpGs with a p value lower than $3.6e-8$ were considered epigenome wide significant based on the theoretical threshold calculated previously (Saffari et al., 2018). Genes associated with epigenome wide significantly differentially methylated CpGs were entered into a GENE2FUNC analysis in “Functional Mapping and Annotation of Genome-Wide Association Studies” (FUMA) online tool (Watanabe, Taskesen, van Bochoven, & Posthuma, 2017) to test whether the genes are overrepresented in FUMA’s predefined gene sets. The GENE2FUNC option in FUMA provides information about likely biological mechanisms that the listed genes are involved in. The tool gives information about diseases, drug targets and expression tissue specificity of each gene. The GENE2FUNC tool also provides information on the level of enrichment for a group of genes in any given biological process. This enrichment analysis is based on existing biological pathway gene sets. One of the libraries of diseases, drug targets and biological pathways, that FUMA refers to, is the Kyoto Encyclopaedia of Genes and Genomes (KEGG).

3.2.4 Differently Methylated Region Analysis

To identify differently methylated regions (DMRs), with 2-30 CpGs within 500bp methylated in the same direction, we used the dmrff function in R (Suderman et al., 2018). CpGs identified in the EWAS with unadjusted p-value of $p \leq 0.05$ were included in the DMR analysis. DMRs with a Bonferroni adjusted p-value of $p \leq 0.05$ are considered statistically significant.

3.2.5 Replication analysis

To replicate the findings from our primary EWAS we performed a least absolute shrinkage and selection operator (LASSO) penalised regression. The training dataset was the DNAm

data from Generation Scotland wave 1 dataset, the within-cohort testing data set was DNAm data from the Generation Scotland wave 2 dataset, and the out-of-cohort testing data set was the DNAm data from the Avon Longitudinal Study of Parent and Children (ALSPAC) data set.

3.2.6 Training dataset

The training dataset was the first wave of the Generation Scotland DNAm data, the pre-processing steps were the same as for the normalised 850k dataset described earlier. We used the 450k dataset as the ALSPAC data set did not have 850k DNAm data available.

Phenotypic and DNAm data was available for 3852 individuals in wave 1. CpG sites were used as the independent variables. SIMD quintile was regressed on age, sex and 10 genetic PCs with the resulting residuals were used as dependent variables in the LASSO regression. SIMD quintile was used rather than rank for the replication analysis as the IMD rank was not available in the out-of-sample testing data set from ALSPAC.

The glmnet function in R was used to run the training model and we applied a 10-fold cross-validation and set the mixing parameter to 1.

3.2.7 Testing Datasets

The within cohort testing dataset was wave 2 of the Generation Scotland 450k DNAm data, the pre-processing steps were the same as for wave 1 450k data and are described in the methods above.

The out-of-cohort testing dataset was the DNAm data from the ALSPAC cohort. The timepoint “Focus on Mothers” clinic visit was chosen because it was the measure of IMD closest in time to the blood draw, from which the DNA methylation data was obtained. The dataset included data on 936 mothers with a mean age of 47.8 years. Normalisation of the Illumina Infinium HumanMethylation450 DNAm data (from ALSPAC) was done using the ‘meffil’ package in R. After normalisation remaining slide effects were regressed out. Meffil was also used to normalise the sample to a reference blood count dataset. The ALSPAC DNAm dataset and the accompanying pre-processing protocol have been described in detail in (Relton et al., 2015).

The SES measure used in the testing data set was the English Indices of Multiple Deprivation (IMD). The IMD is very similar to the SIMD in all but a few ways. One of the ways in which IMD differs from SIMD is that the IMD also includes information about affordability, and thereby access, to housing. The other main difference is that IMD housing domain also includes information on the outdoor environment such as levels of air-pollution and the number of road traffic accidents. IMD and SIMD also differ in the number- and size of data zones. There are 32,844 data zones in the IMD and the average population of an IMD data zone is 1,500 people (The Ministry for Housing, Communities, and Local Government 2015).

The methylation risk scores (MRS) were calculated by multiplying the methylation betas in the testing data with the coefficients obtained in the LASSO regression and summing up the scores for each participant. The predictive accuracy of the resulting MRS was assessed by running a linear regression with MRS as the independent variable and observed derivation score quintile as the dependent variable.

3.3 Results

3.3.1 Demographics

The 7,702 participants from the Generation Scotland dataset, for whom full methylation and phenotypic data was available, had an average age of 49.54 (± 13.61) and a female/male split of 57/43%.

The plot below shows the distribution of SIMD rank in the dataset.

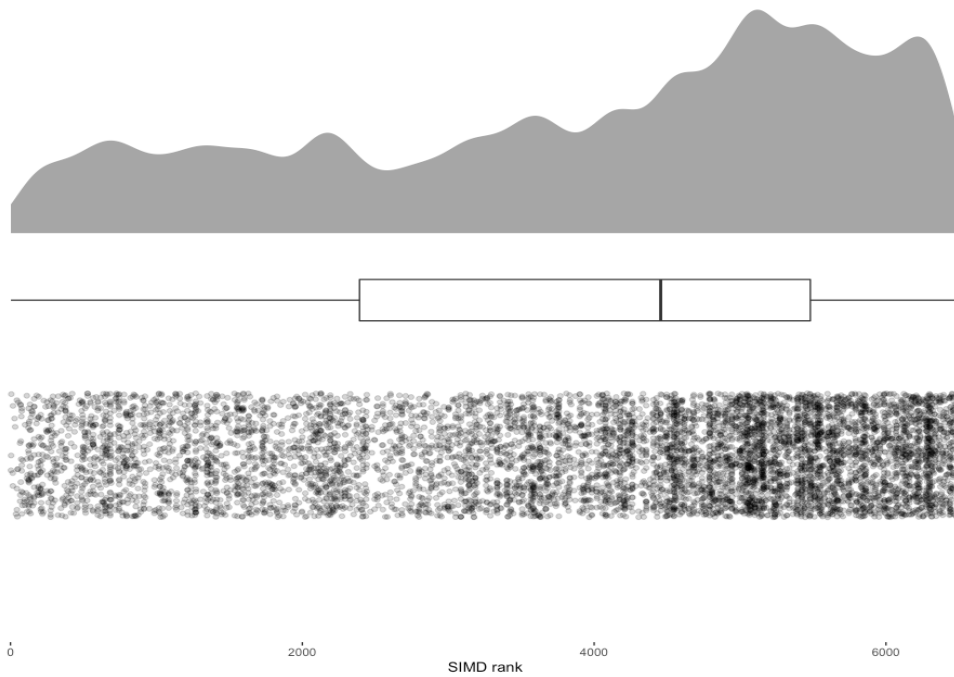


Figure 3.2: Distribution plot of SIMD rank of Generation Scotland participants ($N = 7,702$). An SIMD rank of 1 is the most deprived neighbourhood, an SIMD rank of 6,976 is the least deprived neighbourhood. The vertical axis represents participant frequency, the horizontal axis represents SIMD rank. The top panel is a frequency distribution plot, the middle panel represents a box-plot (minimum, first quartile, median, third quartile, maximum), the bottom panel represents individual datapoints for each participants.

3.3.2 Differentially Methylated Positions

The analysis of differentially methylated positions (DMP) of SIMD rank, controlling for smoking, alcohol, BMI, age, sex, cell counts and PCs, identified 15 CpGs that were significantly associated with SIMD rank ($p < 3.6 * 10^{-8}$). The 15 epigenome-wide significant CpGs are listed in Table 3.1 along with positions, associated genes, p-values and betas.

CpG	Gene	Chromosome	P value	Beta	Standard error
cg03536474	HOXC11	12	1.18E-09	-1.25E-05	2.05E-06
cg07626482	SLC1A5	19	1.57E-09	3.29E-06	5.45E-07
cg13407975	CRLF1	19	3.37E-09	-6.05E-06	1.02E-06
cg17247190		6	3.47E-09	-5.53E-06	9.35E-07
cg21327712	DENND2D	1	3.73E-09	-4.01E-06	6.79E-07
cg20210689	INPP5A	10	4.76E-09	4.68E-06	7.99E-07
cg10575067	ZFP36	19	5.54E-09	-7.68E-06	1.32E-06
cg16800968	HSF4	16	6.30E-09	-9.71E-06	1.67E-06
cg19769182	PRRT2	16	1.47E-08	-6.36E-06	1.12E-06
cg22992427		17	1.56E-08	3.49E-06	6.17E-07
cg00252813	GAPDH	12	2.08E-08	3.97E-06	7.07E-07
cg12828075	INSC	11	2.37E-08	-8.68E-06	1.55E-06
cg05210412	CA4	17	2.49E-08	-8.93E-06	1.60E-06
cg11047325	SOCS3	17	2.94E-08	6.29E-06	1.13E-06
cg19504123	RAB34	17	3.32E-08	-1.02E-05	1.84E-06

Table 3.1: Table of epigenome wide significant CpGs from the SIMD EWAS. CpG name, associated gene, chromosome, p-value, betas, and standard error.

P-values of all CpGs across chromosomes are visualised with a Manhattan plot in Figure 3.3.

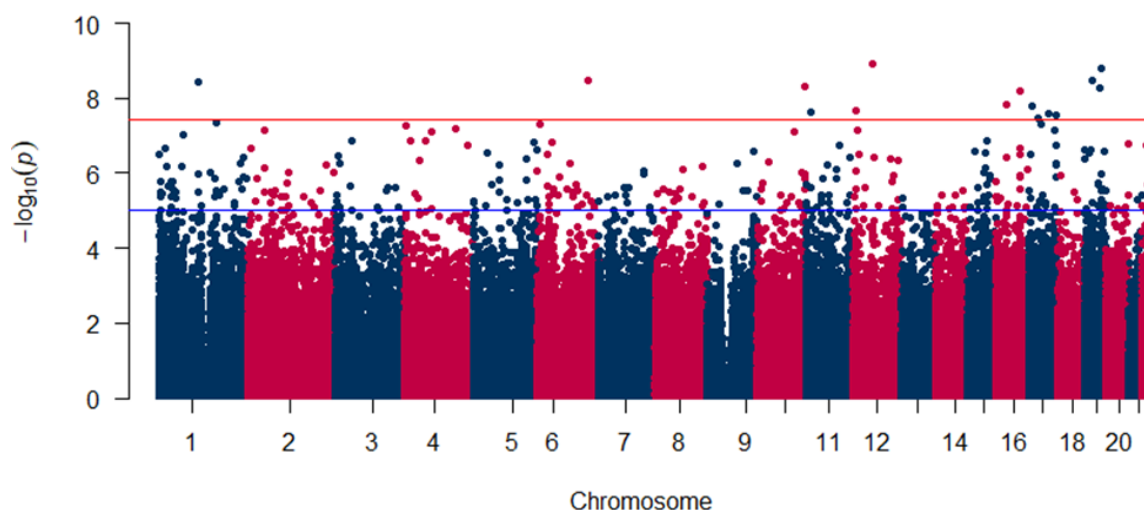


Figure 3.3: Manhattan plot of SIMD EWAS results. The blue line indicates suggestive significance threshold ($p < 1*10^{-5}$). The red line indicates epigenome significance threshold ($p < 1*10^{-8}$).

The 15 CpGs significantly associated with SIMD rank mapped onto 13 unique genes which were entered into the GENE2FUNC tool in FUMA (Watanabe et al., 2017). The

differentially methylated genes were significantly enriched for 14 curated gene sets. The gene sets included the following biological pathways: hypoxia (most commonly caused by asthma, COPD, CVD, smoking, and inflammation), adipogenesis (formation of fat-tissue), mTORC1 signalling (integrates growth factors and nutrients to stimulate growth associated with diabetes, obesity, depression), Tumor Necrosis Factor (TNF, a pro-inflammatory cytokine), prednisolone resistance (a glucocorticoid and anti-inflammatory medication), cyanobacterial LPS antagonist (a natural anti-inflammatory), CD4T cells, spleen monocytes (involved in biological link between inflammation and cancer), upregulation in microglia due to acute inflammation, breast cancer, and KRAS proto-oncogene expression. The gene sets, a heat map of gene expression and the tissue specificity of the differentially methylated genes can be seen in Appendix 1.

To investigate tissue specificity FUMA uses hypergeometric tests to assess whether differentially methylated genes are overrepresented in differentially expressed gene sets (Watanabe, Umićević Mirkov, de Leeuw, van den Heuvel, & Posthuma, 2019). Analysis of tissue specificity of the 13 differentially methylated genes revealed a significant enrichment for genes expressed in the colon, fallopian tube and oesophagus (supplementary materials)

Finally, there was significant upregulation of differentially methylated gene sets associated with brain development in early childhood.

The EWAS catalogue (Battaram et al., 2021) was used to look up the top 15 differentially methylated CpGs in previous literature. Eight of the 15 CpGs had previously been associated with phenotypes at the epigenome wide significance level. The most common phenotype was age, followed by fetal vs adult liver, clear cell renal carcinoma, smoking and inflammatory phenotypes amongst others. A full list of the phenotypes can be seen in Table 3.2.

CpG	EWAS Catalogue Phenotype	Study Author	N	Chr	Position
cg03536474	Age	Mulder et al (2021)	2,338	12	54367460
cg07626482	Clear cell renal carcinoma, Current versus never smoking, C-reactive protein, Gamma-glutamyl transferase, Former versus never smoking	Wozniak et al (2013) Joehanes et al (2019) Nano et al (2017)	129 9,389 1,450	19	47289503
cg13407975	Age, Sex	Mulder et al (2021) Singmann et al (2015)	2,338 1,799	19	18705946
cg17247190	Fetal vs adult liver	Bonder et al. (2014)	195	6	141883437
cg21327712					
cg20210689					
cg10575067					
cg16800968					
cg19769182	Fetal vs adult liver	Bonder et al. (2014)	195	16	29823868
cg22992427					
cg00252813	Age, Primary Sjogren's syndrome, Serum total cholesterol	Mulder et al (2021) Imgenberg-Kreuz et al (2016) Hedman et al. (2017)	2,338 500 4,331	12	6642229
cg12828075	Clear cell renal carcinoma, Age	Wozniak et al (2013) Mulder et al (2021)	129 2,338	11	15136150
cg05210412					
cg11047325					
cg19504123					

Table 3.2: List of differentially methylated CpGs and the phenotypes previously associated with them at the epigenome wide level, according to The EWAS Catalogue, as well as study author, number of participants (N), tissue type, chromosome (Chr.) and position.

The GWAS catalogue (Buniello et al., 2019) was used to look up phenotypes associated with the 13 unique differentially methylated genes. Of the phenotypes there was an overrepresentation of bone health, white and red blood cell counts, height, and respiratory function. A full list of the CpGs, genes and associated phenotypes can be seen in Table 3.3.

CpG	Gene	GWAS Catalogue Phenotype
cg03536474	HOXC11	Spine bone size, systolic blood pressure
cg07626482	SLC1A5	Height
cg13407975	CRLF1	Blood protein levels
cg17247190		
cg21327712	DENND2D	Lung function, COPD
cg20210689	INPP5A	Heel bone mineral density, systolic blood pressure, CVD
cg10575067	ZFP36	Mean corpuscular haemoglobin, Lymphocyte counts
cg16800968	HSF4	Heel bone mineral density, height
cg19769182	PRRT2	Red blood cell count
cg22992427		
cg00252813	GAPDH	Waist-to-hip ratio, cleft lip (tobacco), iris colour
cg12828075	INSC	Heel bone mineral density, breast cancer
cg05210412	CA4	Platelet distribution width, Multiple sclerosis, Monocyte count
cg11047325	SOCS3	Respiratory disease (inflammatory feedback)
cg19504123	RAB34	Height

Table 3.3: A list of phenotypes (according to The GWAS Catalogue) previously associated with genes linked to differentially methylated CpGs

3.3.3 Differentially Methylated Regions

The analysis of differentially methylated regions (DMR) returned 449 regions significantly differentially methylated for SIMD. Table 3.4 shows the top 20 most significant regions.

Chr.	Gene	DMR start	DMR end	Z score	Adj. p-value	Estimate	N (CpG)	Range (bp)
17	ZNF232	5019056	5019989	-9.20595	2.77E-14	-4.06E-06	9	933
6	GNMT	42928337	42928773	-8.80187	1.10E-12	-5.78E-06	12	436
12	GAPDH	6642201	6642953	8.767589	1.49E-12	3.70E-06	7	752
16	ESRP2	68270220	68270882	-8.69912	2.74E-12	-2.65E-06	11	662
19	SLC1A5	47287964	47288263	8.51315	1.38E-11	3.78E-06	7	299
10	ADO	64565336	64565772	8.488718	1.71E-11	3.46E-06	5	436
2		45159504	45160759	-8.47535	1.92E-11	-3.12E-06	13	1255
7	LOC650226	56515510	56516309	-8.37773	4.41E-11	-6.10E-06	9	799
17	ACSF2	48545805	48546620	-8.3606	5.10E-11	-5.53E-06	12	815
17	CHAD	48545805	48546620	-8.3606	5.10E-11	-4.21E-06	12	815
17	SOCS3	76354621	76355202	8.303564	8.26E-11	5.12E-06	5	581
1	G0S2	209848479	209848864	-8.24005	1.41E-10	-8.00E-06	7	385
5	GPX3	150399909	150400531	-8.07328	5.60E-10	-3.42E-06	8	622
4	MARCHF1	165304473	165305815	-7.99128	1.09E-09	-3.29E-06	13	1342
22	HORMAD2	30476098	30476345	-7.98419	1.16E-09	-2.72E-06	9	247
13	EDNRB	78492306	78493203	-7.93458	1.73E-09	-3.37E-06	11	897
8	RGS22	101117952	101118899	-7.93248	1.76E-09	-3.05E-06	12	947
17	RAB34	27044629	27045022	-7.84653	3.50E-09	-4.58E-06	7	393
11	KCNQ1DN	2890473	2891098	-7.83042	3.98E-09	-5.19E-06	18	625
6	TUBB	30689844	30691914	7.811798	4.61E-09	1.27E-06	25	2070

Table 3.4: Differentially methylated regions (DMRs) and their position (chromosome, start and stop), associated gene symbol, adjusted p-value, z-score, estimate, number of CpGs in each DMR, and the number of basepairs in each DMR (Range).

Using the IlluminaHumanMethylationEPICanno.ilm10b2.hg19 (Hansen 2016) 391 unique genes were annotated to the 449 DMRs. A list of the DMRs can be seen in Appendix 2. There is an overlap of five CpGs (cg00252813, cg07626482, cg11047325, cg16800968 and cg19504123) from the 15 epigenome wide significant DMPs with the CpGs included in the DMRs. There is also a considerable overlap between annotated genes for DMP and DMR. In total, 6 out of the 13 DMP annotated genes (GAPDH, HSF4, INPP5A, RAB34, SLC1A5, SOCS3) can be found amongst the 391 unique genes annotated to DMRs.

The 391 unique annotated genes were entered into the GENE2FUNC tool in FUMA (Watanabe et al., 2017) to find associated phenotypes in the GWAS catalogue and to test for enrichment in gene ontology (GO) biological processes and KEGG pathways.

The gene enrichment analysis returned 514 significantly enriched Gene Ontology (GO) pathways in total divided out over biological processes (431), cellular components (39), and molecular components (44). The five most significantly enriched biological processes pathways were animal organ morphogenesis, neurogenesis, cellular component morphogenesis, neuron differentiation, and tissue morphogenesis. The five most significantly enriched cellular component pathways were intrinsic components of plasma membranes, synapses, constituent parts of neurons, synaptic membranes, and constituent parts of synapses. The five most significantly enriched molecular component pathways were sequence specific DNA binding, DNA binding transcription factor activity, regulatory nucleic acid binding, double stranded DNA binding, and sequence specific double stranded DNA binding.

KEGG pathways implicate mostly cellular and tissue scaffolding processes. The enrichment of neuroactive ligand receptor interaction is surprising as it has been implicated in antipsychotic treatment response (Adkins et al., 2012).

GO biological processes point towards early developmental processes which fits with observational studies showing long-lasting and not-easily reversed effects of low SES in childhood. GO cellular processes pointed towards mostly neurological processes, which is surprising given the mostly inflammatory nature of the other phenotypes, except of course for depression and MS. The developmental and neurological pathway enrichment results fit well with the tissue specificity of the differentially expressed genes being significantly enriched for early childhood brain tissue.

GO molecular component pathways are mostly involving DNA binding processes, this could simply be a consequence of dealing with changes to DNAm. DNAm impedes the binding of transcriptional proteins, but also facilitates the binding of methyl-CpG-binding domain proteins to facilitate further epigenetic modifications through the recruitment of, for instance, chromatin. A famous example of this process is the methyl-CpG-binding protein 2 (MeCP2), a loss of which has been seen as a risk-factor for Rett-syndrome (Lyst & Bird, 2015).

A tissue specificity test in FUMA (Watanabe et al., 2019) returned significant enrichment for genes expressed across most tissue types with no differential expression across developmental periods.

A search of the GWAS catalogue (Buniello et al., 2019) returned nine distinct phenotypes, many of which are also present in the DMP analysis. The phenotypes and the contributing gene sets can be found in Table 3.5.

Phenotype Gene Sets	Genes in gene set (N)	Contributing genes (N)	Adjusted P
Systolic blood pressure	793	23	3.99e-3
Sudden cardiac arrest	31	5	8.94e-3
Type 2 diabetes	494	16	1.28e-2
Blood protein levels	1935	37	1.61e-2
LDL cholesterol levels in current drinkers	73	6	2.39e-2
Obesity-related traits	756	19	3.45e-2
Visceral adipose tissue adjusted for BMI	53	5	3.63e-2
Cognitive decline rate in late mild cognitive impairment	122	7	3.63e-2
Major depressive disorder	210	9	3.63e-2

Table 3.5: Phenotypic gene set significantly enriched by the DMR associated genes, as well as the

3.3.4 Lasso prediction of SIMD/IMD quintile

The LASSO penalised regression gave a minimum lambda value resulting in 32 CpG sites and accompanying weights. The CpGs and weights can be seen in appendix 3. These weights were applied to the DNAm m-value at each CpG in the internal testing sample (Generation Scotland wave 2) and the external testing sample (ALSPAC).

For the internal replication the SIMD derived methylome risk scores (MRS) significantly predicted the observed SIMD quintile in Generation Scotland wave 2 accounting for 1.5% of the variance ($\beta = 3.00$, $p = 1.85 \times 10^{-13}$, $R^2 = 0.015$).

A linear model comparing the MRS-estimated IMD scores in ALSPAC with the observed IMD scores in ALSPAC showed no significant prediction ($\beta = -0.46$, $p = 0.065$, $R^2 = 0.0025$).

3.4 Discussion

The aim of this study was to explore epigenome wide differences in DNA methylation associated with area-based deprivation specifically the Scottish Index of Multiple Deprivation. The single largest methylation dataset to date was used to conduct the EWAS and report 15 differentially methylated points and 449 differentially methylated regions associated with deprivation.

The 15 DMPs were associated with area-based deprivation at the epigenome wide level after controlling for two measures of smoking behaviour, units of alcohol consumed and measured BMI.

A search of the GWAS catalogue for phenotypes associated with the differentially methylated genes from the DMP analysis returned a mix of bone health (heel bone mineral density, red blood cell count), height, inflammatory processes (lymphocyte count, monocyte count, Chronic Obstructive Pulmonary Disorder), cardiovascular and respiratory diseases. Phenotypes associated with differentially methylated genes from the DMR analysis paint a similar picture; cardiovascular traits (blood pressure, sudden cardiac arrest), inflammatory traits (blood protein levels, type 2 diabetes), nutritional phenotypes (cholesterol, visceral adipose tissue, obesity related traits) but also major depressive disorder and mild cognitive impairment. These findings point to an environmentally driven association between area-based deprivation and a host of chronic physical and mental health conditions, in particular of an inflammatory origin over and above the effect of smoking tobacco, drinking alcohol and increased BMI. The present results are congruent with previous well-established findings on associations between low SES and inflammatory diseases, cardiovascular disease, respiratory diseases and depression (Milaniak & Jaffee, 2019; Rosengren et al., 2019; Sahni et al., 2017).

The enrichment of differentially methylated genes for both DMP and DMR were not highly differentiated in tissue types. But the DMP enrichment analysis showed differential enrichment for gene sets preferentially upregulated during brain development in early childhood. This finding fits well with research showing that an individual's SES during early childhood is associated with difference in disease risk well into adulthood and old age, regardless of present SES status (Hertzman, 1999; Warren, 2009). In future studies it would be interesting to further investigate whether differences in DNAm patterns are associated with SES during early childhood over and above that of adult SES.

DMP associated genes were primarily enriched for inflammatory processes, cancers, mTORC1 signalling, hypoxia, and adipogenesis. mTORC1 signalling is involved with

regulation of homeostasis and metabolism and has previously been implicated in obesity related complications. It has also been associated with immunoregulations and inflammatory processes as well as neurodegenerative diseases and major depressive disorders (Dazert & Hall, 2011). Genes associated with tumor necrosis factor was also significantly enriched, it is a pro-inflammatory cytokine and has previously been found to be involved in insulin resistance, diabetes and Alzheimer's disease (Bellenguez et al., 2022; Sethi & Hotamisligil, 2021).

Differentially methylated genes from the DMP analysis were significantly enriched for a gene set previously associated with prednisolone resistance. Prednisolone is a corticosteroid analogous to the body's glucocorticoid and used to treat inflammatory diseases. Prednisolone resistance is most common in COPD and severe asthma cases and is thought to arise from extreme levels of oxidative stress interfering with the expression certain histone deacetylases thereby hindering the anti-inflammatory actions of prednisolone.

Hypoxia is involved in both respiratory- (e.g. COPD), and cardiovascular diseases (congestive heart failure) (J. W. Lee, Ko, Ju, & Eltzschig, 2019). All of these biological processes were also associated with differentially methylated genes identified in the DMR analysis.

DMR associated genes were primarily enriched for organ morphogenesis, neurogenesis, neuron differentiation, focal adhesion, neuroactive ligand receptor interaction and regulation of actin cytoskeleton. Regulation of actin cytoskeleton has been associated with immunological processes in general and in particular with respiratory and vascular diseases (Tang & Gerlach, 2017; Wickramarachchi, Theofilopoulos, & Kono, 2010).

Previous epigenome wide association studies have identified phenotypes associated with seven out of the 15 DMPs. Those phenotypes were age, inflammatory markers, cancer, cholesterol, former vs. never smoker, gamma-glutamyl transferase and fetal vs. adult liver. None of the studies listed in Table 3.2 included the Generation Scotland sample used in the present EWAS.

Age was associated with four of the 15 DMP despite being included as a covariate in the EWAS model. It is perhaps to be expected since age has some of the most pervasive effects on DNAm. However, we would not expect there to be any measurement error in the recording of chronological age that could introduce residual confounding. The age associated differences in DNAm after covarying for chronological age may be indicative of accelerated biological ageing. Several previous studies have established DNAm signatures related to

biological ageing, coined accelerated epigenetic ageing, and shown associations with common adverse health outcomes (Lu et al. 2019). A recent study using a DNAm based predictor of lifespan called GrimAge, found a significant association between educational attainment accelerated GrimAge, or in other words increased risk of mortality (Hillary et al., 2021). These findings mirror what we know from epidemiological studies; that low SES is associated with decreased life expectancy. It is however worth pointing out that the study (Mulder et al. 2021) associating age with the abovementioned CpGs found roughly half of the 450,000 CpGs sampled to be associated with age which could be due to inflation whereby the level of statistical significance is overestimated.

The many differentially methylated genes, regions and enriched pathways pointing to inflammatory diseases and mental health problems, in particular major depressive disorder fits well with previous literature pointing to systemic chronic inflammation being linked to mental health problems. The present findings present the possibility that inflammatory processes are part of the mechanisms behind the association between low SES and depression. However, the biological mechanisms linking inflammatory diseases and depression are also not properly understood.

3.4.1 Strengths and limitations

Limitations of the current study include cross sectional nature of the data and possible residual confounding of obesity related traits and smoking. Due to the cross-sectional nature of the DNAm data collected we cannot draw inferences about direction of causality of the associations found. It is, however, unlikely that changes in DNAm lead directly to low SES – although it is possible that they could lead to changes in disease liability which in turn might contribute to the risk of living in low SES. It seems likely that what we are observing is an archive of environmental signals associated with low SES. The results could however, also include disease-associations as well as genetic associations with low SES that are conferred through changes in DNAm associated with low SES (Barbu et al. 2021). Despite controlling for BMI, smoking, and alcohol in the EWAS model there still seems to be an association with obesity and smoking related traits and some of the differentially methylated genes identified. BMI, alcohol consumption and smoking have pervasive effects on the epigenome. In this study we used a self-report measure of alcohol consumption however, recent evidence from genome wide association studies suggest that participants with a higher disease burden in

particular tend to underreport their alcohol consumption (Xue et al., 2020). Smoking have some of the strongest known effects on DNAm and it is possible that even with two smoking variables as covariates we still have a slight residual confounding of smoking. One possibility for future studies is to use body fat percentage as a covariate instead of BMI and a DNAm derived smoking score as a covariate instead of smoking status and pack years. In particular the AHRR related CpG (cg05575921) has been found to be highly associated with smoking. Adding individual levels of DNAm levels as a covariate may remove residual confounding from smoking.

Prescription medications were not included as covariates in the present study. As a result it is possible that some of the signal from, for example, type 2 diabetes and depression is due to treatment of the disease changing DNA methylation at CpGs or genes previously associated with said disease. If such a confounding is present it is not necessarily a problem since the findings still reflect the biological impact of living in a deprived neighbourhood. However, it is important to keep in mind before drawing any causal conclusions.

The replication analysis in the form of a least absolute shrinkage and selection operator (LASSO) regression returned a significant prediction of SIMD quintile within the Generations Scotland cohort, using wave 1 of DNAm data as a training dataset and wave 2 as a testing dataset. To my knowledge, this is the first time that neighbourhood deprivation levels have been predicted using DNAm derived scores. It is however, worth noting that the predicted SIMD scores only explained 1.5% of the variance in the observed SIMD scores. The coefficients derived from wave 1 Generation Scotland data could not provide a significant prediction of IMD in the external prediction analysis of ALSPAC data. The lack of replication could be due to differences between the two cohorts. The ALSPAC FOM DNAm is an all-female dataset with an average age of 47.8 years. Furthermore, All the participants were recruited as new mothers over a short period of time, their age range is relatively narrow. The Generation Scotland wave 1 sample has an average age of 48.1 years with a SD of 13.7. Age has a great impact on DNAm patterns and this difference in age range could impact the replicability of results.

The reason for the lack of external validation may also be found in the deprivation measures themselves. SIMD and IMD are slightly different from one another both in terms of weighting of modules, the exact measures included and the number of people in each data zone. But most importantly the deprivation measures are relative, and ranked. If one cohort is more deprived than the other thereby skewing the predicted scores. Because the deprivation is ranked it is not possible to compare the two cohorts directly to find out whether one is

significantly more deprived than the other. The difference in prediction results could be down to differences in pre-processing of the data despite the fact that similar chips and pre-processing were used in both ALSPAC. No matter the reason for the lack of significant prediction, it emphasises the need for investigating how stratified DNAm levels are between physical localities, time points, age groups and social groups.

The use of commercially available pathway enrichment analysis tools, such as FUMA, does come with methodological limitations. The tools are developed for gene expression arrays and as such they do not account for the fact that some genes have a much higher number of CpGs associated with them, than others. This imbalance becomes even more pronounced when genes previously associated with common diseases are more likely to be represented by CpGs on the sequencing chips. Although pathway enrichment analyses are powerful tools for investigating underlying biological mechanisms, it is important to keep these limitations in mind when interpreting the results (Harper, Peters, and Gamble 2021).

The present study draws its strength from being the single largest EWAS of area-based deprivation. This has provided the opportunity investigate DNAm differences associated with low SES with a previously unobtainable precision as well as using a richer phenotype that better captures the many facets of deprivation. Finally, our findings are congruent with previous research showing associations between low SES and adverse health conditions such as inflammatory diseases, decreased life expectancy, neurodegenerative disease and major depressive disorder.

3.4.2 Future directions

Air-pollution has been found to aggravate inflammatory disease symptoms through a SOCS3 pathway in rat liver (Long et al., 2019). The SOCS3 gene (previously associated with inflammatory respiratory disease) was differently methylated both in the present DMP and DMR analysis. Because of the residual confounding of smoking behaviour observed in the DMP associated phenotypes, it is difficult to determine whether the difference in methylation at the SOCS3 gene is due to smoking behaviour or air-pollution. Future investigations may want to interrogate the persistence of inflammatory respiratory diseases after adequately correcting for smoking. Unfortunately, air-pollution data is not available in the Generation Scotland cohort. Instead one could use the available measure of urbanicity as a proxy for air-pollution. It would be interesting to see if adding urbanicity as a covariate would remove the

signal from respiratory diseases in general and the SOCS3 gene in particular. However, the SIMD working group does caution against using SIMD from rural areas as the data zones there cover a larger geographical area and therefore include more varied levels of deprivation than the urban data zones. Urban vs. rural living has also been associated with other mental health disorder such as schizophrenia (Vassos et al. 2012, Mortensen 2000), it would be interesting to exploration possible differences in the depression-SES association across the urban-rural gradient.

Finally future studies of DNAm differences associated with neighbourhood deprivation could compare results to individual level SES factors. For example, a comparison of DNAm differences associated with educational attainment and DNAm differences associated with SIMD may give an indication of overlapping or separate biological impacts for the two phenotypes.

In summary we performed an EWAS of area-based deprivation in more than 7,700 individuals from the Generation Scotland cohort. We found 15 CpGs with differences in DNA methylation reaching epigenome wide significance and 449 differentially methylated regions. Phenotypes associated with differentially methylated genes and enriched gene-pathways suggest inflammatory processes, cardiovascular disease, bone health and respiratory diseases are associated with socioeconomic status. These findings are congruent with, and extend previous research as discussed.

4 DNA Methylation Signatures of Inflammatory Disease and Depression

4.1 Introduction

Depression is the number one cause of years lived with disability world-wide (WHO 2017). As such there is a clear need for more effective treatments for depression. A recent meta-analysis reported that while antidepressant medication are effective for the majority of patients (Cipriani et al., 2018), up to 20% of depressed individuals are resistant to therapy (Mrazek, Hornberger, Altar, & Degtiar, 2014). The differences in effective treatment for all patients may result from causal heterogeneity. Emerging evidence suggests that people with elevated levels of systemic inflammation exhibit a weakened treatment response and may even represent a subgroup of depression with a different molecular aetiology (Gardner & Boles, 2011).

Emerging evidence suggests that depression is associated with inflammatory diseases and increased concentrations of inflammatory markers in both the blood (Berk et al., 2013) and cerebrospinal fluid (Lindqvist et al., 2009). Recent GWAS studies have shown that several of the genetic loci associated with depression are known to be involved in immune-related processes (Naomi R Wray et al., 2018) and a recent DNA-methylome wide study of genetic risk of depression found associations with immune response (Shen et al. 2021). The evidence therefore seems to point towards a possible causal role for inflammation in depression, although the direction of causation remains unclear (Kappelmann, Lewis, Dantzer, Jones, & Khandaker, 2018). Changes to DNA methylation (DNAm) have also been implicated in both depression and in multiple inflammatory diseases (Story Jovanova et al., 2018),(McDermott et al., 2016) suggesting that this could give an insight into the molecular connection between inflammation and depression.

DNAm is a type of epigenetic modification which regulates the transcription and splicing of genes. During methylation, a methyl group is added to the DNA strand at cytosine-phosphate-guanine (CpG) sites. Importantly levels of DNAm changes with environmental exposures in a non-random manner and has been documented to change in a dose dependent relationship (McCartney et al., 2018). These attributes make it a prime

candidate for discovering molecular signatures of environmentally mediated processes. Previously our group have successfully used DNAm signatures to predict depression status, to investigate novel biological underpinnings of anti-depressant use, and to explore structural brain changes associated with inflammation (Miruna C Barbu et al., 2019); (M C Barbu et al., 2020; C. Green et al., 2021).

A recent study involving 200 participants, (Crawford et al., 2018) found 6 Differentially Methylated Regions (DMRs) (LTB4R, HOXC4, TRIM39, DNAJC17, PNPLA2, and MEF2A) but no Differentially Methylated Points (DMPs) associated with the interaction of depression and inflammatory disease. The study showed that DNAm could be a useful tool for exploring underlying biological mechanisms of inflammatory depression. In a similar vein DNAm signatures, of otherwise volatile, systemic inflammation markers have proved a valuable method in investigating brain anatomical changes related to depression (C. Green et al., 2021).

The first chapter of this thesis described an association between low SES and depression and quantified that association. The analysis of the biological impacts of living in low SES in the previous chapter revealed differences in methylation at epigenetic and genetic loci previously associated with inflammatory markers and diseases, cardiovascular phenotypes, and major depressive disorder. These findings fit well with previous literature and highlights the possibility that inflammation is the biological mechanism underlying the association between low SES and depression. These findings justify further investigation of the mechanisms underlying the association between depression and systemic chronic inflammation.

The aim of the present study is to leverage the environmentally dependent plasticity of DNAm and use DNAm differences as environmental molecular signatures, to explore the biological mechanisms underpinning inflammatory depression. To do this we will perform an epigenome wide association study (EWAS), with a 40-fold increase in sample size to previous studies, into the interaction of inflammatory disease and depression status whilst covarying for known confounding factors such as antidepressant medication use, smoking and alcohol use.

4.2 Methods

Phenotypic and DNA methylation data were obtained from the family-based population cohort study Generation Scotland comprising more than 24,000 individuals (between 2006 and 2011) from across Scotland, with both baseline cross sectional (B. H. Smith et al., 2013) and longitudinal (Navrady et al., 2018) data. The present study included N=8,700 individuals for whom data was available on DNAm, inflammatory disease, depression, smoking behaviour, BMI, and alcohol consumption. Phenotypic data and the blood draw occurred at the same baseline visit between 2006 and 2011.

GS:SFHS received ethical approval from NHS Tayside Research Ethics Committee (REC reference number 05/S1401/89) and has Research Tissue Bank Status (reference: 15/ES/0040). Written informed consent was obtained from all participants.

4.2.1 Phenotypes

Inflammatory disease was scored from baseline self-report questionnaire data on the presence versus absence of any chronic inflammatory condition (which included diabetes, osteoarthritis, rheumatoid arthritis, COPD, or asthma). Depression was diagnosed by a trained researcher at the baseline visit using the SCID (Spitzer, Gibbon, Skodol, Williams, & First, 2002) after an interview with the SCID screening questions for mood disorder.

Use of anti-depressant medication was assessed through a self-report questionnaire. Body mass index (BMI) was calculated using height (cm) and weight (kg) measurements recorded by clinical staff at baseline visit. Alcohol consumption was recorded in a self-report questionnaire at baseline. In the present study we used the number of units consumed in the week prior to the baseline visit. We used 2 measures of smoking: smoking status and pack years. All three measures were recorded with a self-report questionnaire at the baseline visit. For smoking status participants indicated whether they currently smoke, used to smoke but stopped within the last 12 months, used to smoke but quit more than 12 months ago or never smoked. Pack years were calculated by multiplying the number of cigarettes smoked per day with the number of years the participant has smoked.

The smoking and alcohol consumption phenotypes underwent some processing to retrieve possible data loss due to answering questionnaires in an unintended way. This pre-processing of the smoking and alcohol consumptions phenotypes was developed after the

previous chapter resulting in a larger number of participants being included in the present chapter.

In the Generation Scotland dataset there are three alcohol related variables:

- *Drinking status*, "Have you ever had an alcoholic drink" (1 - Yes, currently drink, 2 - Yes, but stopped within past 12 months, 3 - Yes, but stopped more than 12 months ago, 4 - No, never drank).
- *Units*, "During the past week, please record how many units of alcohol you have had" (units).
- *Usual*, "How does this compare-what you usually drink in a week?" (1 - More, 2 - Same, 3 - Less)

Out of 23,548 entries there are 1,854 NAs in the "*units*" variable. 1391 of those are for participants who answered 4 (no never drank) or 3 (Yes, but stopped more than 12 months ago) in "*drinking status*". Note the rest of the "*units*" NAs are probably to be found among the 509 NAs in "*drinking status*". In fact, out of the 1,910 participants that answered 4 or 3 (stopped or never drank) in "*drinking status*" only 496 answered 0 for units, which would have been the intended way of filling out the survey. 23 participants answered 1 or more (max = 8) to the "*units*" question ("During the past week, please record how many units of alcohol you have had") despite not having touched an alcoholic drink for at least 12 months.

Finally, the "*usual*" variable ("How does this compare-what you usually drink in a week?") showed a similar pattern of answers: 1,706 of the 1,910 participants who never drank or stopped drinking (3 or 4 on "*drinking status*") were NAs for the "*usual*" variable.

As a result, participants with NAs in "*units*" that answered 3 or 4 in "*drinking status*" were included in the analysis as a 0 for "*units*" and a 2 ("same") for "*usual*" giving up to another 1706 participants worth of data to analyse. Out of those 1706 participants with extra data points in the units "variable" 600 participants had full DNAm and phenotypic data to be included in the present EWAS.

4.2.2 DNA methylation

Epigenome-wide DNAm was obtained from 9,873 individuals using Illumina HumanMethylationEPIC (850k) BeadChip array. Data collection was performed in two

waves, while pre-processing and quality control was done in one batch. Data collection and pre processing has already been described in chapter 3 section 3.2.2 as well as in Barbu et al. (2021). The final dataset included corrected M-values at 713,522 loci. There was phenotypic and DNAm data for 8,700 participants.

The collection and pre processing of the DNAm data was carried out by another team at The University of Edinburgh, Dr Rosie Walker in particular did much of the data processing.

4.2.3 Statistical methods

The open access statistical software programme R version 3.5.1 was used for all statistical analyses and data visualisations.

4.2.3.1 Epigenome Wide Association Study

Epigenome wide associations were estimated using linear regression models, computing a linear regression for each CpG probe on a chromosome-by-chromosome basis. To investigate DNAm differences in depression cases with a chronic inflammatory disease as opposed to depression cases without it, each model regressed DNAm levels over an inflammatory disease and depression interaction term, correcting for antidepressant medication use, smoking (pack years and status), BMI, alcohol consumption, and the wave in which the DNAm data was collected (see model below). The interaction term (*inflammatory disease*depression*) was made the independent variable and the main effects (*depression status, inflammatory disease*) were included. This approach was adopted to enable inferences to be made about DNAm associated with the cooccurrence of depression and inflammatory diseases *over and above* the occurrence of depression alone or inflammatory diseases alone. In other words, the focus of the EWAS was to investigate whether there were unique differences to the patterns of DNAm in people who present with both depression and an inflammatory disease as opposed to people who only have one of the diagnoses or neither of them. Anti-depressant medication is very broadly (approaching universally) prescribed to patients with depression and at the same time almost exclusively prescribed to treat depressions. As such it was included as a covariate to remove any group differences in DNAm patterns solely due to differences in prescriptions of said drug.

*CpG ~ inflammatory disease * depression status + depression status + inflammatory disease + antidepressant medication use + smoking status + pack years + BMI + alcohol consumption + data collection wave.*

Any CpGs with a p value lower than 3.6×10^{-8} were considered epigenome wide significant based on the threshold calculated previously by Saffari et al (Saffari et al., 2018). Any CpGs with a p value lower than 1×10^{-5} were considered to have a suggestive association.

4.2.3.2 Functional Mapping and Annotation of Genome-Wide Association Studies

Genes associated with differentially methylated CpGs at the discovery threshold were entered into a GENE2FUNC analysis in “Functional Mapping and Annotation of Genome-Wide Association Studies” (FUMA) online tool (Watanabe et al., 2017) to test whether the genes are overrepresented in FUMA’s predefined gene sets.

4.2.3.3 Differently Methylated Region Analysis

To identify differently methylated regions (DMRs), with 2-30 CpGs within 500bp methylated in the same direction, we used the dmrff function in R (Suderman et al., 2018). CpGs identified in the EWAS with unadjusted p-value of $p \leq 0.05$ were included in the DMR analysis. DMRs with a Bonferroni adjusted p-value of $p \leq 0.05$ are considered statistically significant.

4.3 Results

4.3.1 Inflammatory Disease and Depression

A logistic regression of inflammatory disease and depression controlling for age, sex, BMI, smoking, alcohol consumption, and antidepressant medication, revealed a significant association between the independent and dependent variables ($\beta = 0.205$, std error = 0.071, $p = 0.0039$). The results indicate that in the Generation Scotland cohort, the odds of having depression are 22.8% (95% CI [1.06, 1.41]) higher for people with one or more inflammatory diseases.

4.3.2 Differentially Methylated Positions

The analysis of differentially methylated positions (DMP) of the interaction of inflammatory disease and depression, controlling for antidepressant medication use, smoking, BMI, and alcohol consumption, did not identify any sites at the epigenome wide significance level.

The analysis did however identify eight differentially methylated CpGs at the discovery threshold ($p < 1 \times 10^{-5}$). These 8 discovery threshold CpGs and their p value, estimate, chromosome and associated gene can be seen in Table 4.1.

CpG	Closest Gene	Chromosome	Estimate	P-value	Standard Error
cg08697545		16	-0.0502	2.54E-06	0.0107
cg04864199	SHF	15	0.0626	2.82E-06	0.0134
cg08959747	POMT2	14	-0.0636	4.04E-06	0.0138
cg00748566		5	-0.0358	6.11E-06	0.0079
cg10010898	AXIN1	16	-0.0430	8.55E-06	0.0097
cg14928734	PTMA	2	0.0195	8.72E-06	0.0044
cg12092051		15	-0.0498	9.21E-06	0.0112
cg20936226	MYLK	3	-0.0381	9.84E-06	0.0086

Table 4.1: Top CpGs identified in EWAS of inflammatory disease and depression status interaction.

P values of all CpGs included in the analysis are visualised in a Manhattan plot in Figure 4.1.

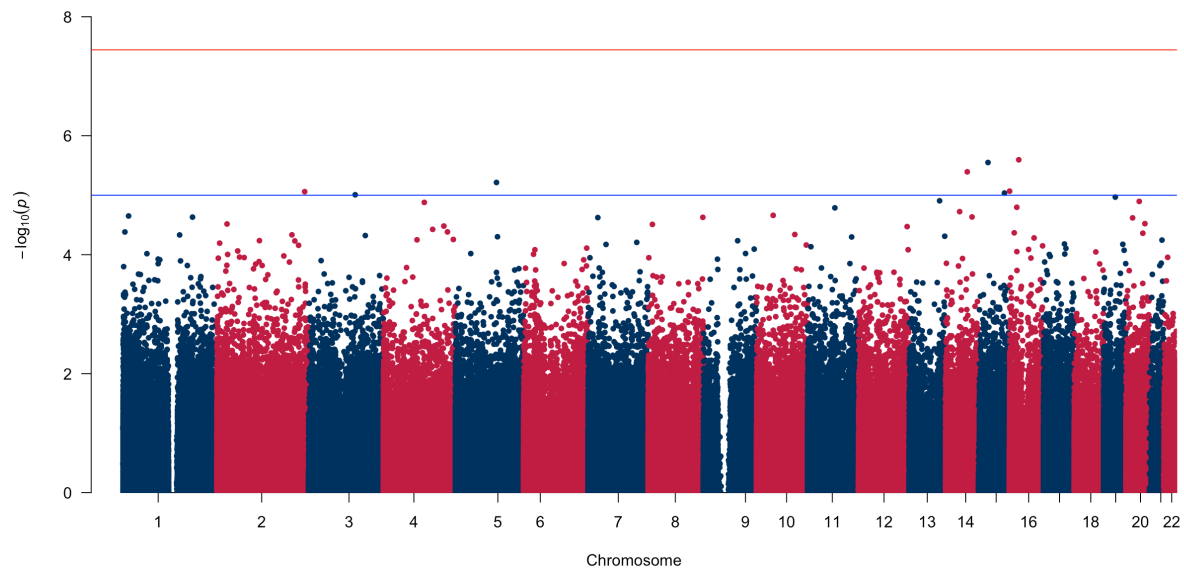


Figure 4.1: Manhattan plot of inflammatory disease and depression status interaction. Alternating blue and red panels to indicate different chromosomes, blue line = discovery significance threshold ($p < 1 * 10^{-5}$), redline = epigenome wide significance threshold ($p < 1 * 10^{-8}$).

The EWAS catalogue (Battram et al., 2021) was used to look up phenotypes previously associated with the discovery threshold DMPs. These included age, air-pollution, toxin exposures and alcohol consumption. The results of the EWAS catalogue search can be seen in Table 4.2.

CpG	Phenotype	Study Author	N	Chr	Position
cg04864199	Age	Mulder et al. (2021)	2,338	15	45493327
	Prenatal phthalate exposure	Soloman et al. (2017)	332		
cg10010898	Age	Mulder et al. (2021)	2,338	16	339598
	Air pollution exposure	Panni et al. (2016)	2,956		
cg14928734	Age	Mulder et al. (2021)	2,338	2	232572399
	Alcohol consumption	Liu et al. (2018)	9,643		
cg20936226	Age	Mulder et al (2021)	2,338	3	123382403

Table 4.2: CpGs and previously associated phenotypes, as well as study author, number of participants (N), tissue type, chromosome (Chr.) and position.

The GWAS catalogue (Buniello et al., 2019) was used to look up phenotypes associated with the five differentially methylated genes. Amongst others, phenotypes include renal- and hepatic health markers, blood protein levels, obesity related traits, and cardiovascular disease. A full list of the differentially methylated genes and associated phenotypes can be seen in Table 4.3.

Gene	Phenotype
SHF	Uric acid levels
	Response to alcohol
	Bisphosphonate-associated atypical femoral fracture
	Creatine levels
	Estimated glomerular filtration rate
POMT2	Response to paliperidone in schizophrenia
	Protein quantitative trait loci (liver)
	Red blood cell count
AXIN1	Adult body size
	Blood glucose levels
	Body mass index
	Cardiovascular disease
	Height
	Haemoglobin concentration
	Iron status biomarkers (ferritin levels)
	Male-pattern baldness
	Mean corpuscular volume
	Metastatic colorectal cancer survival in treatment with chemotherapy plus biologics
	Predicted visceral adipose tissue
	Red blood cell count
	Red cell distribution width
	Systolic blood pressure
	Bone mineral density
	Waist circumference
PTMA	White blood cell count
	RANTES levels
	Lymphocyte counts
	Monocyte percentage of white cells
	Neutrophil count
	White blood cell count
	Mean corpuscular haemoglobin
MYLK	Adolescent idiopathic scoliosis
	Educational attainment
	Cognitive empathy
	General cognitive ability
	Core binding factor acute myeloid leukemia

Table 4.3: DMP associated genes and previously associated phenotypes.

The differentially methylated genes, above the discovery threshold, were entered into the GENE2FUNC tool in FUMA. The 5 genes were significantly enriched for 1 gene set, assessing the association between physico-chemical features and toxicity associated pathways (Antczak, Ortega, Chipman, & Falciani, 2010). There was no significantly differentially expressed genes in the queried tissue specificity gene sets.

4.3.3 Differentially Methylated Regions

The analysis of differentially methylated regions returned 7 regions that were significantly differentially methylated for the inflammatory disease and depression interaction.

Using the IlluminaHumanMethylationEPICanno.ilm10b2.hg19 catalogue (Hansen 2017) 5 unique genes were annotated to the 7 DMRs. The 7 regions can be seen in Table 4.4.

Chr.	Gene	N (CpG)	DMR start	DMR end	Range (bp)	z-value	Adj. P-value	Estimate
20	GNAS	20	57427274	57427977	703	6.9830	2.07E-06	0.0264
19	B3GNT3	5	17905332	17905718	386	5.7928	4.97E-03	0.0124
12		3	2031262	2031546	284	5.7338	7.04E-03	0.0437
1	STMN1	7	26233332	26233565	233	5.5656	1.87E-02	0.0439
12	EID3	3	104697532	104697983	451	5.5425	2.14E-02	0.0510
12	TXNRD1	2	104697532	104697983	364	5.5425	2.14E-02	0.0641
17		5	406137	406501	703	5.4285	4.08E-02	0.0474

Table 4.4: DMRs associated with the inflammatory disease and depression interaction. The table includes position (chromosome, start and stop), associated gene symbol, adjusted p-value, z-score, estimate, number of CpGs in each DMR, and the number of base pairs in each DMR (Range).

The 5 differentially methylated genes from the DMR analysis were entered into the GENE2FUNC tool from FUMA. Two of the 5 genes were significantly enriched for one hallmark gene set relating to glycolysis and gluconeogenesis (Liberzon et al., 2015). There was no significantly differentially expressed genes in the queried tissue specificity gene sets. A search of the GWAS catalogue returned phenotypes previously associated with all of the 5 DMR associated genes. Phenotypes previously associated with the differentially methylated genes can be seen in Table 4.5.

Gene	Phenotype
GNAS	Blood glucose levels Blood protein levels Blood urea nitrogen levels Body size at age 10 Chromosomal aberration frequency (chromatid type) in genotoxic compound exposure Chronic lymphocytic leukemia Colorectal cancer Core binding factor acute myeloid leukemia Creatinine levels Cystatin C levels Diastolic blood pressure DNA methylation (variation) Event-related brain oscillations Glomerular filtration rate HDL cholesterol levels Height Hyperuricemia Insulin-like growth factor 1 levels Mean arterial pressure Platelet attributes Post bronchodilator FEV1/FVC ratio
B3GNT3	Serum cancer antigen 19.9 levels Strep throat Tumor biomarkers
STMN1	Hepatocyte growth factor levels Waist-to-hip ratio adjusted for BMI
TXNRD1	Stem cell growth factor beta levels Emphysema annual change measurement in smokers (percent low attenuation area) Hypertrophic cardiomyopathy (sarcomere positive) Obesity-related traits Logical memory (delayed recall) in normal cognition Protein quantitative trait loci (liver) Height Toenail selenium levels Response to serotonin reuptake inhibitors in major depressive disorder
EID3	Unipolar depression, response to selective serotonin reuptake inhibitor

Table 4.5: DMR related genes and previously associated phenotypes.

4.3.4 Replication

4.3.4.1 DMP overlap

Five out of the six CpGs that Crawford et al. (2018) identified in their candidate CpG analysis of inflammatory disease and depression interaction were present in this study. Out of those 5 CpGs one of them reached nominal significance in the present EWAS analysis (cg11411220, $p = 0.039$).

4.3.4.2 DMR overlap

None of the differentially methylated genes identified in (Crawford et al., 2018) DMR analysis were nominally significant in the present DMR analysis.

4.3.4.3 SIMD EWAS overlap

None of the 15 CpGs identified as epigenome wide significantly differentially methylated in the SIMD EWAS analysis reached nominal significance in the EWAS of inflammatory disease and depression.

4.4 Discussion

In the present study we aimed to explore the molecular background of the link between chronic inflammatory disease and increased risk of depression. To that end we ran an epigenome wide association of the interaction of chronic inflammatory disease and depression.

The Epigenome Wide Association Study did not reveal any differentially methylated CpGs at epigenome wide significance. However, it did identify 8 differentially methylated CpGs reaching discovery threshold significance. Phenotypes previously associated with these discovery DMPs include age, alcohol consumption, prenatal phthalate exposure, and air pollution. Gene annotation revealed five unique genes located near the eight discovery DMPs. These genes have previously been associated with blood protein levels, inflammatory markers, BMI and waist circumference, blood glucose levels, bone mineral levels, and creatine and urea levels.

The DMR analysis revealed seven significantly differentially methylated regions, which encompassed 5 unique genes. Phenotypes previously linked to the DMR associated genes were mostly associated with inflammatory markers, kidney function, liver function, blood pressure, obesity related traits, growth factor (Insulin like and stem cell), and treatment response to SSRIs.

Pathway enrichment analysis for both DMP and DMR associated genes returned few results, but most notably two of the DMR related genes were significantly enriched for a hallmark gene set associated with glycogenesis.

Previous research (Crawford et al. 2018) into DNAm signatures of inflammatory disease and depression did not identify any DMPs at either an epigenome wide significance threshold or a discovery threshold. None of the top ten CpGs identified in previous research overlap with the discovery threshold DMPs in the present study. Similarly, none of the 45 CpGs making up the seven DMRs identified in the present study, overlapped with any of the 30 CpGs making up the six DMRs identified by Crawford et al. (2018). It is worth noting that Crawford et al's analysis included 200 participants whilst the present study included 8,700 participants. The difference in power between the two studies may explain the lack of replication.

The previous association of age with four out of the eight DMPs is interesting as age was corrected for in the pre-processing of the DNA methylation data. The association could be attributable to increased epigenetic ageing, as both inflammation and depression have previously been linked to increased epigenetic ageing (L. K. M. Han et al., 2018; Stevenson et al., 2018). It could also be a time-to-event effect with a build-up of inflammation over the life course. A recent study into DNAm derived c-reactive protein (CRP) scores showed an increase in CRP with advancing age (Stevenson et al., 2020). As mentioned in chapter 3 it is worth pointing out that the study, in which the CpGs in question were linked to age, found an association between age and DNAm levels in close to 50% of probes (Mulder et al. 2021). The high number of significantly differentially methylated points may suggest inflation of the statistic.

Phthalate exposure has previously been associated with asthma (Franken et al., 2017), BMI, waist circumference (Deierlein et al., 2016) and chronic inflammation (Bai et al., 2017). Preclinical studies have linked phthalate exposure to downregulation of

BDNF expression and hippocampal changes (C. A. Smith & Holahan, 2014) both of which are known to change with depression. Increased levels of phthalate metabolites have been associated with increased depressive symptoms (K.-S. Lee et al., 2018). Finally, as mentioned in the introduction, exposure to phthalates has been linked to both dysregulation of BDNF (Rowdhwil & Chen, 2018) and aberrant HPA axis functioning (Hlisníková et al., 2021). To the best of our knowledge this study is the first to show a possible connection between phthalate exposure and inflammatory depression using DNAm signatures. We suggest future research investigates the relationship between environmental phthalate exposure, chronic inflammation, and depression.

Air pollution, previously associated with methylation levels at cg10010898, is known to increase the risk of chronic inflammation and inflammatory diseases.

(Anderson, Thundiyil, & Stolbach, 2012; K.-

J. Chuang, Chan, Su, Lee, & Tang, 2007; Hajat et al., 2015). Air pollution levels have also been associated with an increased risk of psychiatric disorders, including depression (Antonsen et al., 2020; Lim et al., 2012). It has been suggested that small particulate matter can cause neurotoxicity through inflammation and oxidative stress thereby contributing to neurodevelopmental and neurodegenerative disorders (Block & Calderón-Garcidueñas, 2009; Costa et al., 2017). In the light of these previous findings, it is perhaps expected that the present study finds an association between inflammatory disease and depression, and differences in DNAm at CpGs previously associated with exposure to air pollution. Based on these results we suggest future research focuses on investigating the links between air pollution exposure, chronic inflammation, and depression, as well as possible confounders between inflammatory depression and these environmental stressors.

None of the studies previously linking the discovery DMPs from this EWAS to other phenotypes, included the Generation Scotland sample.

The fact that one of the discovery DMPs have previously been associated with alcohol consumption could be due to residual confounding. Alcohol consumption is known to have pervasive effects on DNAm (Liu et al., 2018), similar to the effect of smoking (Joehanes et al., 2016). While we included two variables in our analysis to control for smoking, we only included one variable (units consumed in the previous week) to control for alcohol consumption. The rate at which DNAm changes associated with smoking recover, to DNAm patterns of non-smokers, is dose-dependent (McCartney et al., 2018). If DNAm changes due to alcohol consumption follow a similar timeline it might be worth including a

variable accounting for the length of time since individuals stopped drinking, as non-drinkers may still have the DNAm changes associated with alcohol consumption if they recently quit. There is also a possibility that the association of inflammatory depression with DNAm differences at a CpG previously associated with alcohol consumption is due to shared biological pathways between the phenotypes. In the study that associated cg14928734 with alcohol consumption it was reported that many of the genetic loci identified in the EWAS were associated with immune related functions (Liu et al., 2018).

The DMR gene enrichment analysis revealed two genes being significantly enriched for a hallmark gene set associated with glycolysis and glycogenesis. Both are metabolic processes, glycolysis is involved in converting glucose into ATP, whereas gluconeogenesis is the process of generating glucose from non-carbohydrate substrates. Interestingly for this analysis gluconeogenesis has been found to be disrupted by insulin resistance which in turn can be triggered through cortisol induced inflammation (Menkin, 1943). Aberrant glycolysis, gluconeogenesis and insulin resistance are all symptoms of type II diabetes. Other diabetes related traits (blood glucose levels and obesity related traits) appeared as associated phenotypes of both DMP and DMR related genes.

Phenotypes associated with diabetes and insulin resistance were anticipated in our results since diabetes was one of the inflammatory diseases used to calculate the inflammatory disease and depression interaction. Of note, insulin resistance has been proposed as a shared mechanism behind depression and type 2 diabetes (Lyra E Silva et al., 2019) and metformin, a common drug used to treat diabetes has shown to have some antidepressant effects (M. Guo et al., 2014).

The pathway enrichment of gluconeogenesis also fits the HPA theory of depression described in the general discussion. Glucocorticoids, such as cortisol, curbs inflammation but are also involved in both gluconeogenesis and glycolysis. It is possible that the pathway enrichment results reflect an overactivated glucocorticoid system which would not only affect metabolic processes but also the risk of depression through negative feedback system in the hippocampus associated with decreased neurogenesis and, over time, hippocampal damage.

On a more unexpected note, and linking back to a phenotype previously associated with one of the discovery DMPs, air pollution has repeatedly been associated with increased risk of diabetes (Y. Li, Xu, Shan, Teng, & Han, 2019).

Response to SSRIs was associated with two of the DMR related genes. This finding is intriguing given previous literature on differences in treatment response to common antidepressant medications in inflammatory depression (Arteaga-Henríquez et al., 2019; Lindqvist et al., 2017). A GWAS of the concentrations of citalopram and ecitalopram in blood from patients treated with citalopram or ecitalopram, identified genome wide significant associations (Ji et al., 2014). Concentrations of citalopram/ecitalopram and their metabolites in blood gives an indication of the level of biotransformation taking place and thereby also the individuals response to a given drug.

We controlled for antidepressant medication use in the EWAS model suggesting that the resulting DMRs are not medication vs. no medication artefacts. The DMRs in question both showed increased methylation associated with inflammatory disease and depression, this increased level of methylation could repress gene transcription thereby influencing biotransfer of SSRIs and ultimately impact treatment response. The findings tentatively raise the hypothesis that systemic chronic inflammation could change treatment response through changes in methylation at genetic loci involved in drug metabolism. The present findings highlight the importance of a systematic investigation of antidepressant treatment response in depressed individuals with co-occurring chronic inflammatory diseases.

Liver and kidney functioning seem to be especially prevalent phenotypes in our results. Liver and kidney related phenotypes associated with the DMP genes include, uric acid levels, creatine levels, estimated glomerular filtration rate and protein quantitative trait loci (liver). Liver and kidney related phenotypes associated with the DMR genes include blood urea nitrogen levels, creatine levels, cystatin c levels, glomerular filtration rate, hyperuricemia, protein quantitative trait loci (liver). Chronic inflammation has previously been associated with creatine and cystatin C levels (Kaysen, 2009) which in turn are associated with increased risk of cardiovascular disease (Muslimovic, Tulumovic, Hasanspahic, Hamzic-Mehmedbasic, & Temimovi, 2015). Cardiovascular disease is also one of the phenotypes associated with both DMP and DMR genes in the present study. A common factor between inflammation, depression, cardiovascular disease and, liver and kidney functioning is air pollution (Kim, Min, Seo, & Min, 2019). We propose future research should include renal and hepatic diseases when investigating chronic inflammatory diseases and depression. Furthermore, based on the underlying common factor of air pollution and air pollution appearing as a

directly related phenotype to one of the DMPs, we suggest that future research focus on investigating the effects of air pollution on inflammatory depression.

Kidney and liver functioning is involved in the clearing of toxins and in metabolic processes. These findings are congruent with the phthalate and air pollution exposure previously being associated with the DMPs in this study, and with metabolic pathways revealed in the gene enrichment pathway analyses. Renal and hepatic health does not immediately spring to mind when considering biological processes underlying inflammatory depression, but it is reasonable to assume that xenobiotics impacting immune function and mental health also have hepatorenal consequences.

Finally, the abundance of inflammatory markers among the phenotypes associated with differentially methylated genes in the DMP and the DMR analysis were expected and serves as an important validation check.

Similar to the previous EWAS there is a possibility that some of the biological signal in the results is originating from treatments associated with the diseases we are investigating (depression, diabetes, asthma, COPD, osteoarthritis, and rheumatoid arthritis) rather than the biological mechanisms underlying the diseases themselves. However, the focus of the current EWAS was the interaction of inflammatory disease and depression over and above inflammatory diseases or depression on their own. Using the interaction term as the independent variable and the main effects as covariates increases the possibility of interpreting the results as specific to the interaction of the inflammatory diseases and depression. Consequently, any signal due to anti-inflammatory medications should be controlled for by the inclusion of inflammatory disease as a covariate, unless individuals with both inflammatory disease and depression have a different treatment profile. These are intriguing questions and future research should aim to control for medication use in their analyses. It is not possible to rule out that some of the differences in DNAm can be due to differences in the use of anti-inflammatory medications. Anti-inflammatory medications are prescribed in a much less specific manner (often used as pain relief for a wide range of conditions). As such it was decided not to include them as a covariate in this analysis, it is however an area of much needed research, and it would be interesting to explore differences in DNAm patterns based on anti-inflammatory medication in the future.

The present DMP and DMR results did not replicate any of the results from the results from the previous study looking at DNAm differences associated with inflammatory disease and depression (Crawford et al. 2018). There are some methodological differences between the two analyses, for instance the present study included anti-depressant medication use as a covariate which the previous study did not. Despite still being underpowered, the present study also had more than 40-fold the sample size of the previous study which might explain some of the differences between results. However, given the lack of between cohort replication in the previous chapter of this thesis it may be of importance to investigate the cross-sample robustness of DNAm patterns in general.

The main limitation of the current study is its sample size. Despite being the largest study of its kind to date, a larger sample is needed to detect epigenome wide differences in DNA methylation associated with the interaction of inflammatory disease and depression. It is estimated that roughly four-fold the sample size needed to detect a main effect, is needed to detect an interaction effect (Fleisz 1968). The only study, to date, to report differentially methylated CpGs ($N = 3$) associated with depression at an epigenome wide significance threshold ($p = 1.57 \times 10^{-8}$) is Jovanova et al (2018). To perform this study, they used a cohort of 7,948 participants. Consequently, a rough estimate of the sample size needed to detect a differentially methylated CpG for the interaction of inflammatory disease and depression, with a statistical significance lower than $p = 1 \times 10^{-8}$, is a little over 31,000 individuals. There is no such single DNAm dataset available and even aggregating cohorts fall well short of that number. However, this serves highlight the need for much larger epigenetic epidemiology cohorts in the future.

Future studies may also benefit from including protein-based measures of chronic inflammation. A new and promising method is to use a more stable DNAm score for the otherwise volatile CRP measure that is normally used to indicate chronic inflammation (Stevenson et al. 2020; Green et al. 2021; Conole et al., 2020).

On a larger scale it might be beneficial for the cohort to adjust their questionnaires of smoking and alcohol consumption, or the scoring off them, going forward. As explained in the methods section, it seems likely that the current questionnaires lead to information loss. To more accurately control for smoking behaviour in DNAm studies it may also be helpful to develop a smoking behaviour questionnaire that more accurately reflects the temporal components of changes to DNAm due to smoking. In the current dataset there is a variable containing information whether an individual is a current versus former smoker. Previous

literature has shown that pervasive smoking related changes to DNAm remain for years after smoking cessation. Going forward it may make more sense to have a variable coding how long ago a person has regularly been smoking.

In summary, we tested associations between differences in DNA methylation and the interaction of inflammatory disease and depression, to shed light on the biological underpinnings of inflammatory depression.

We discovered no epigenome wide significant DMPs, but significant DMRs and discovery threshold DMPs, associated with inflammatory disease and depression, were identified. The DMPs and the DMRs have previously been linked to air pollution, age, phthalate exposure, inflammatory markers, renal and hepatic related phenotypes, and diabetes related phenotypes.

Findings from the present study highlight possible biological pathways between systemic inflammation and depression, as well as identifying novel potential environmental risks for inflammatory depression.

5 Discussion

5.1 Introduction

The aim of this thesis was to investigate the relationship between socioeconomic status and depression and to explore the biological mechanisms underlying their association.

Depression is one of the most debilitating diseases in the world with a relatively high prevalence and rate of recurrence it has a significant emotional, social, and economic cost for the individual, their families and society. Treatments for depression are effective for most patients but tangible targets for preventative measures are lacking. Low socioeconomic status (SES), a measure of educational attainment, income, and employment type, has consistently been associated with a heavy disease burden. A higher prevalence of depression has also been observed in low SES settings which may offer an avenue for preventative measures. Unfortunately, the magnitude of the increased risk of depression associated with low SES is not well documented with regards to changes at the molecular, cellular, and systems levels.

The aims of this thesis were therefore:

- 1) To establish a cross-cultural and cross-developmental estimate of the risk of depression associated with low SES.
- 2) To explore the biological impact of living in low SES or multiple deprivation.
- 3) To look for possible mechanisms connecting the strongest biological signal from living in multiple deprivation, i.e. inflammatory diseases, and depression.

The thesis includes three separate investigations: a systematic review of the association between socioeconomic status and depression, an epigenome wide association study (EWAS) of area-based deprivation, and an EWAS of the interaction of inflammatory disease and depression.

The aim of the systematic review and meta-analysis was to quantify the risk of depression associated with low SES based on the largest possible evidence base, to provide risk estimates across cultures and developmental time-periods, and to look for differences in risk of depression between SES sub-categories and country-level Gross Domestic Product (GDP) categories.

The aim and rationale of the area-based deprivation EWAS was to use the non-static and environmentally dependent nature of DNAm levels to examine the long-term biological

impact of living in a low SES setting. A measure of area-based deprivation, the Scottish Index of Multiple Deprivation (SIMD) was used as an indicator of SES as it more accurately represents the lived environment of individuals at a population level, than measures solely based on income, educational attainment, or occupational status. DNAm represses or increases transcription of a gene at particular sites throughout the genome, and therefore is a potential candidate mechanism for mediating associations between low SES and MDD. However, due to the intricacies of studying naturally occurring and highly complex traits, the inferences in this thesis only focussed on the ability of DNAm levels to change with environmental experiences and exposures, as a form of a biological archive or “log-book”, and not on its ability to cause direct changes to gene expression.

The aim of the EWAS of the interaction between inflammatory disease and depression was to investigate possible biological mechanisms connecting inflammatory diseases and depression. The majority of differentially methylated CpGs and genes in the SIMD EWAS have previously been associated with immune processes and inflammatory markers and diseases, suggesting that living in deprived neighbourhoods impacts an individual’s immune system and inflammatory pathways. Previous literature has linked systemic chronic inflammation with a higher risk of depression, suggesting that this could be one of the biological pathways linking low SES with a higher risk of depression. To investigate how depression with an inflammatory aetiology is different from depression alone or inflammatory disease alone an EWAS of inflammatory disease and depression was performed.

Each empirical chapter in this thesis includes a discussion section specific to said chapter therefore in this general discussion chapter I will briefly discuss the findings of each empirical chapter in turn detailing how the findings fit together and relate to the existing literature on SES, depression, and inflammation. Strengths and limitations of the methods and models used in this thesis will also be discussed. Finally, I will provide suggestions for improvement on the methodological approaches used in this thesis and I will give recommendations on how to use the findings from this thesis to further investigate the association between SES and depression.

5.2 Summary of main findings

5.2.1 Low socioeconomic status, in particular low income, is associated with increased risk of depression across cultures.

The literature search for the systematic review identified just under 8,000 publications. The titles and abstracts of this literature were screened, leaving 521 publications to be full text screened. A total of 140 studies were included in the systematic review, with 112 of those entering the meta-analysis. Together the meta-analysed studies included 329,005 participants, of which 48,158 were clinically relevant depression cases. The participants came from 47 different countries spanning 6 continents. This makes it the geographically widest spanning meta-analysis of SES and depression.

The meta-analyses revealed an overwhelming degree of evidence of low SES being associated with a higher risk of depression and this difference in risk was observed across all sub-domains of SES. Out of the 112 studies that were meta-analysed only three individual studies showed odds in the opposite direction.

Of the three SES sub-domains, the strongest risk of depression was associated with low income. It is possible that low income has a more direct effect on health, including mental health, than education, employment, or a composite SES score. This could be because low education or unemployment often lead to a lower income. Education may provide other beneficial effects such as increased cognitive reserve or be indicative of a more stable childhood. Similarly, being in employment may provide stronger self-confidence, an increased sense of purpose and maybe a wider support network through colleagues. However, income level is potentially more directly related to psychological stress and the opportunity to engage in healthy behaviours such as healthy eating and exercising but also whether or not you have the temporal- and mental capital to deal with minor illness and ailments early on. Similarly, in many countries income level determines whether an individual can access even basic medical care, and in most countries whether you can access specialist care such as counselling or psychological therapy. An individual's level of income strongly determines their quality of housing both in terms of the actual residence (insulation, dampness and mould, lead piping, presence of asbestos, etc) but also the area and the infrastructure and amenities that come with it. Over time, or during sensitive developmental time-periods, these small differences in behaviours and physical surroundings may compound and eventually

create the population level differences in overall health, including risk of depression, observed in this study.

The meta-analysis of ranked or ordinal data showed a dose dependent relationship between low, medium, and high SES and the risk of having depression. These findings suggest that even incremental changes to an individual's level of SES could influence depression risk. However, the meta-analyses were only performed in cross sectional data so any conclusions on longitudinal changes cannot be made from these results. The findings from the included longitudinal studies were summarised in a narrative review and are largely congruent with the results of the meta-analyses; low SES, compared to high SES, is associated with an increased risk of depression.

The results from the meta-analyses had a relatively high level of heterogeneity. Sensitivity analyses found no significant differences in depression risk associated with low SES between income economic groups (high vs. low-income economies). This finding is intriguing as it is the first time, to my knowledge, that the risk of depression associated with low SES has been demonstrated to be comparable across cultures. It is also worth pointing out that the most recent systematic review and meta-analysis of the association between depression and neighbourhood SES show similar levels of heterogeneity despite more narrow inclusion criteria (Richardson et al. 2015). It is possible the high level of heterogeneity between studies is due to differences in recruitment setting. A substantial proportion of the studies recruited participants from hospital settings. These patients often have severe or cooccurring diseases which may encompass different aetiologies than milder single case morbidities, and therefore different risk estimates for low versus high SES. Clinical populations also tend to have a higher prevalence of depression, than the general population, which again can introduce difference in the risk estimates between studies. Another source of heterogeneity could be the differences in depression screening tools used in the included studies. It is likely that each diagnostic tool classifies depressed vs. not depressed in slightly different ways leading to small differences in the prevalence of people diagnosed as depressed with low or high SES. In the systematic review I attempted to limit this source of bias by including only studies comparing levels of SES between people without depression and people with clinically relevant depression, as opposed to comparisons across the whole spectrum of depressive symptoms.

The findings of low SES, in particular low income, being associated with increased risk of depression across cultures could be an important piece of information in the

translation of observational epidemiological research into developing strategies to prevent depression. The present thesis shows that low income is associated with a 69% increase in the risk of having depression, the highest of any of the SES measures. This finding provides a tangible focus for preventative approaches to depression mitigation. Most studies of the association between SES and depression are limited to an observational design. As a result, they are more susceptible to bias than experimental and randomised interventional experiments. However, the strength of observational studies is to provide a framework on which to generate hypotheses for interventional experiments.

Income is potentially susceptible to intervention, within the context of an individual, or cluster randomised trial - were the experimental group are given more monetary resources. As mentioned in the discussion of the systematic review and meta-analysis, one such approach, universal basic income, is already being trialled in several countries and pilot schemes have already shown great promise. However, as with any trial it is difficult to determine what a meaningful outcome measure might look like, let alone to quantify it. Differences in health may not become appear or at least be measurable in the short to medium term. As mentioned in the general introduction there's mounting evidence suggesting low SES in childhood can have profound effects on adult health (Warren et al. 2009). If that is the case, interventions may have to last a full generation instead of months to years, to observe a meaningful difference in disorder risk. To perform a more targeted preventative trial aiming to minimise or mitigate the association between low SES on depression we need to know exactly how low SES may impact the individual's biology and ultimately, health. To get a better understanding of the biological impact of living in a deprived area I performed an epigenome wide association study of area-based deprivation.

5.2.2 Area-based deprivation is associated with widespread differences in DNAm patterns, most of which are associated with inflammatory processes.

The EWAS of the area-based measure of deprivation, SIMD, returned 15 epigenome wide differentially methylated points and 449 differentially methylated regions. In the EWAS model I included two measures of smoking, a measure of BMI and a measure of alcohol consumption as covariates to get an impression of the biological impact of deprivation over

and above the lifestyle factors that are traditionally known to cause health problems associated with low SES.

A measure of area-based multiple deprivation was used in this analysis in an attempt to more accurately capture the lived experience of an individual. Given that the systematic review and meta-analysis found the strongest association between low income and depression it would have been interesting to compare the results of an EWAS of SIMD rank with the results of an EWAS on personal income. This was not done due to time-constraints and the prioritisation of other thesis aims. However, comparing differences in DNAm patterns of income and SIMD would be a great candidate for future work.

Most of the associated phenotypes were derived from GWAS studies and studies on transcriptional pathways. In other words the differentially methylated CpGs are associated with genes which in turn have been associated with phenotypes in previous literature. The field of studying trait specific DNAm differences in humans is relatively young and as such there is not a lot of previous literature directly linking CpG probes to specific traits. A search of the EWAS catalogue did however return previously associated phenotypes for seven out of the 15 epigenome wide DMPs. The phenotypes were; age, clear cell renal carcinoma, differences between fetal and adult liver tissue. Gamma-glutamyl transferase (GGT), smoking, C-reactive protein, primary Sjogren's Syndrome, and serum total cholesterol. As mentioned in the chapter discussion, the appearance of age as an associated phenotype was unexpected as this was covaried for in the EWAS model. A search through a previous study linking the four DMPs with differences in age (Mulder et al., 2021), revealed that the difference in methylation at each CpGs was the mirror opposite to the pattern found in the present study. Bearing in mind that an increase in SIMD rank means a decrease in deprivation, the present findings could point to increased deprivation being associated with epigenetic accelerated ageing. This is of course something that future studies should test specifically.

The previous study (Mulder et al., 2021) investigated blood samples collected at multiple time points from birth to late adolescence, a developmental time-period that has previously been shown to be particularly vulnerable to detrimental effects of deprivation. Psychological stress and physical hardship have also been shown to accelerate the onset of puberty (James-Todd, Tehranifar, Rich-Edwards, Titievsky, & Terry, 2010). Epigenetic ageing has recently become a research area of increasing interest. As mentioned in the chapter discussion lower educational attainment has been associated with increased epigenetic mortality risk as measured by the GRIMage clock (Hillary et al 2021). Childhood

abuse, neighbourhood deprivation, and financial disadvantage has been associated with epigenetic deviations from chronological age using the Horvath and Hannum epigenetic clocks (Marini et al., 2020). There is evidence that early childhood SES predicts epigenetic aging of immune cells better than current SES status as adults (Austin et al., 2018) suggesting there may be particularly sensitive periods in childhood where environmental impacts have long-lasting effects on DNAm. To my knowledge this is the first time we see epigenome wide significant differences in whole blood DNAm linking low SES to increased epigenetic ageing.

The fetal vs. adult liver phenotypes, clear cell renal carcinoma, and GGT, an enzyme associated with liver disease, all point to differences in liver functioning being associated with deprivation. High levels of GGT can also be indicative of diabetes, congestive heart failure or liver cancer, including clear cell renal carcinoma. Higher risk of liver cancers in general have been associated with low SES (Joshi et al. 2008) and with increased alcohol consumption (Turati et al., 2014). Alcohol consumption has a pervasive impact on DNAm levels right across the epigenome (Lohoff et al. 2021) and it is possible that the alcohol consumption covariate did not adequately control for that impact. The alcohol consumption measure used was a questionnaire item on the number of units consumed in the last week. I chose this variable rather than a questionnaire item on the number of units usually consumed, as the ‘units in the last week’ measure was consistently higher, and I suspect having to remember a usual number of alcohol units consumed could result in a more biased estimate. Studies on changes in DNAm over time show that environmental impacts can have relatively long-lasting effects (Cardenas et al., 2017; Colicino et al., 2021). In which case it might be more useful to either use a longitudinal measure of alcohol consumption or create a DNAm based alcohol predictor to use as a covariate. A longitudinal measure would very likely be difficult to obtain in most studies and may still be biased. As mentioned earlier, clinical populations tend to underreport their alcohol consumption (Xue et al., 2020) thereby skewing the distribution of results. On the other hand, a DNAm derived alcohol consumption score could remove any and all effects that are associated with increased alcohol consumption but not necessarily directly linked to it. A similar situation is likely to be the case for the smoking variables associated with the cg07626482 probe in chapter 3. As discussed previously the inclusion of a DNAm derived smoking score, DNAm levels at the AHRR associated CpG, or smoking questionnaires designed with temporal dynamics of a DNAm smoking patterns in mind, may remove the residual confounding seen in the present analysis.

The final four phenotypes; CRP, primary Sjogren's syndrome (an autoimmune disorder), and serum total cholesterol are congruent with the phenotypes associated with the differentially methylated genes, namely of a metabolic and inflammatory nature.

The phenotypes associated with the differentially methylated genes from the DMP and the DMR analyses can be divided into five rough groups: inflammatory processes, cardiovascular traits, bone health, metabolic traits, and mental health. Most of the phenotypes were related to the immune system, inflammatory processes and inflammatory diseases. These findings are congruent with previous literature, as mentioned in the general introduction low SES has previously been associated with inflammatory diseases. In this regard the SIMD EWAS findings provide a molecular biological confirmation of the existing evidence of health problems associated with low SES. Cardiovascular disease (CVD) has also been associated with low SES in the past, and both inflammatory states and metabolic traits are risk factors of CVD. The metabolic traits associated with the differentially methylated genes could be due to residual confounding of obesity. BMI was included in the EWAS model to account for obesity related traits. However, BMI may not adequately capture the effects of obesity, let alone differences in fat distribution. This could be the reason why 'visceral adipose tissue (VAT) adjusted for BMI', referring to fat tissue around the inner organs, showed up in the results. There have been arguments for VAT being more important risk factor for health problems, including cardiovascular disease and dementia, than BMI (Cereda, Sansone, Meola, & Malavazos, 2007). In DNAm studies it might be better to include body fat percentage instead of BMI to control for obesity related traits. In hindsight it makes sense that when we investigate changes on the molecular scale it becomes even more important to choose covariates that are as biologically close the phenotype of interest as possible. At a broader perspective, when designing and implementing measures to document or prevent so-called lifestyle disease it may be more advantageous to focus on the ratio, and distribution of fat rather than the dichotomous obese/normal weight.

One of the more surprising phenotypes to consistently show up in both DMP and DMR results was bone health, bone growth, and traits related to height. It is possible that these phenotypes reflect the impact of differences diet and nutritional health, maybe even from a very early developmental stage. Previous candidate DNAm studies of SES have found links between low SES and differentially methylated genes previously associated with bone health (Stringhini et al. 2015, McGuinness et al. 2012). These effects even seem to go back centuries with skeletons from class-segregated cemeteries showing bone health deficits

associated with lower social class (Miszkiwicz & Cooke, 2019; Ng & Chin, 2021). A review by (Ng & Chin, 2021) suggest psychological stress and systemic inflammation can negatively impact bone health through the endocrinological system. They argue that increased levels of glucocorticoids can have a detrimental effect on osteoclastogenesis and osteoblastogenesis (bone growth) thereby leading to bone-loss over time. The association with increased levels of glucocorticoids fit well with the glycolysis and glycogenesis results from the inflammatory disease and depression EWAS, and further support the notion of low SES being linked to increased levels of stress hormones; a well-known risk factor of depression.

The respiratory disease phenotypes could of course be linked smoking and stem from inadequate correction in the model. However, indoor, and outdoor air pollution could also be involved. A recent systematic review found housing-related risk factors associated with respiratory diseases (Wimalasena, Chang-Richards, Wang, & Dirks, 2021). In particular, damp and mouldy air as a result of poor ventilation was found to be a risk factor of respiratory disease. Outdoor air-pollution has also been linked to respiratory disease. The SOCS3 gene that was differentially methylated in the DMP and the DMR analysis has been associated with inflammatory respiratory disease specifically in preclinical experiments mimicking natural air-pollution by using particulate matter with a diameter $\leq 2.5 \mu\text{m}$ (PM_{2.5}) as a physiological stressor (Long et al. 2020).

The appearance of Multiple sclerosis (MS) is interesting, as it is unclear whether SES is a risk factor for this neurological disorder. A systematic review found inconclusive evidence of a link between the SES and MS (Goulden, Ibrahim, & Wolfson, 2015). However, risk factors for MS include type 1 diabetes, inflammatory bowel disease and smoking, obesity and even lower vitamin D exposure. Perhaps there is a biological overlap in the risk factors of MS and low SES. Similarly, a recent cohort study found that low SES was a risk factor for worsened disease progression and increased levels of disability of MS patients (Harding et al., 2019).

The DMR analysis returned 449 region mapping onto 391 unique genes. Five pathways from the Kyoto Encyclopaedia of Genes and Genomes (KEGG) were significantly enriched, mostly implicating cellular and tissue scaffolding processes. The enrichment of the KEGG pathway neuroactive ligand receptor interaction is unexpected, as it has been implicated in antipsychotic treatment response (Adkins et al., 2012). The gene enrichment analysis returned 514 significantly enriched Gene Ontology (GO) pathways. The most significantly enriched pathways were morphogenesis (organ, tissue, and cellular)

neurogenesis, synapses, constituent parts of neurons and synapses, and DNA binding processes. Together these findings may point towards early stages of development as particularly vulnerable to the impact of living in a deprived area, an impact that may be long-lived. The enrichment of brain related pathways is somewhat surprising as the analysis was done with blood-derived DNAm data on the impact of deprivation in general. However, these findings align with the multiple sclerosis and depression phenotypes appearing in the DMP and DMR results. The concentration of brain-related processes and tissue types in the pathway enrichment results are also reflected in the significant upregulation of differentially methylated genes associated with brain development in early childhood. These findings may indicate that multiple deprivation has a particularly strong impact on brain development and brain health, including the risk of depression. These findings no doubt warrant further scrutiny but in time they may help to point to highlight the importance of making sure young children are not experiencing the hardships of deprivation as they may be particularly vulnerable to it.

5.2.3 Exploring biological mechanisms connecting inflammatory diseases and depression

To investigate possible biological mechanisms underlying the cooccurrence of the strongest biological signal from the SIMD EWAS, namely inflammatory disease, and depression, I performed an EWAS of inflammatory disease and depression. The independent variable was the interaction of inflammatory disease and depression to allow findings to be specific to inflammatory depression over and above inflammatory disease or depression on their own. Using an interaction does, however, impact the sample size needed. Roughly, the sample size needed to find an effect of a given size is four times as large for an interaction term as the one needed to find an effect of a similar size in one of the main terms.

(Story Jovanova et al., 2018) found 2 epigenome-wide significant CpGs associated with depressive symptoms in a sample of 7,948 individuals. This suggests that an interaction study of depression and inflammatory disease would need a sample size around 30,000 individuals.

No epigenome wide DMPs were identified in the inflammatory disease and depression EWAS. However, eight CpGs reached a p-value discovery threshold ($p < 1 \times 10^{-5}$) and seven significant DMRs were identified. There was no overlap in results with a previous study (Crawford et al 2018) into the interaction of inflammatory disease and depression. The

lack of replication may be due to lack of power to detect meaningful effects. The present analysis used the largest sample to date, with 8,700 participants and was underpowered to detect interaction term effects, Crawford et al (2018) included 200 participants and found no significant effects.

Phenotypes previously associated with DNAm differences at the discovery threshold CpGs included age, air pollution, phthalate exposure, and alcohol consumption. Age was also associated with DMPs from the EWAS of area-based deprivation and once again this result may reflect accelerated epigenetic ageing being associated with the cooccurrence of inflammatory disease and depression. Exposure to external environmental toxins air pollution and phthalate have both been previously associated with increased risk of several inflammatory diseases and an increased risk of depression (van den Bosch et al. 2019). Finally previous research has linked differences in DNAm at one of the discovery DMPs with alcohol consumption. This result is likely due to residual confounding despite having included a phenotypic measure of alcohol consumption as a covariate in the model.

The eight discovery DMPs were located near five unique genes that had previously been associated with blood protein levels, inflammatory markers, BMI and waist circumference, blood glucose levels, bone mineral levels, and creatine and urea levels (indicative of liver and kidney functioning), treatment response to antipsychotics as well as cognitive ability and educational attainment. The differential methylation of genes previously associated with educational attainment and cognitive ability, observed in individuals with inflammatory disease and depression, is intriguing. The analysis was run to test for biological mechanisms underlying a relationship between inflammatory disease and depression, because an exploration of the biological impact of multiple deprivation primarily returned inflammatory disease related phenotypes. It is possible that this result could reflect an association between low SES and increased risk of depression with an inflammatory aetiology.

The DMRs were situated close to five unique genes that had previously been associated with: inflammatory markers, kidney function, liver function, blood pressure, obesity related traits, growth factor (Insulin like and stem cell), and treatment response to SSRIs.

The kidney and liver function phenotypes could be due to residual confounding from alcohol consumption, similar to the results from the discovery threshold DMP. However, these results could be linked to some of the other phenotypes associated with DMPs and

DMRs. For instance, exposures to environmental toxins such as phthalates and air pollution impacts kidney and liver functioning (Cai, Fan, Ye, Rao, & Li, 2021; B. Guo et al., 2022; Kuźma, Małyszko, Bachórzewska-Gajewska, Kralisz, & Dobrzycki, 2021). The liver and kidney phenotypes from the DMR gene enrichment analysis also fit well with the gene enrichment pathway results of glycolysis and glycogenesis. These processes are associated with metabolic diseases like diabetes (one of the diseases making up the inflammatory disease variable in the EWAS model) but they are also linked to glucocorticoid levels and functioning of the HPA axis in general.

Overall, the findings from this analysis possibly point to distinct biological signatures of inflammatory disease and depression over and above any of the two alone. The DMR related genes associated with treatment response to SSRIs is a particularly intriguing result, since we included antidepressants as covariate. These results may indicate that independent of previous medication status, people with depression *and* an inflammatory disease respond differently to certain antidepressants compared to people with only depression *or* an inflammatory disease.

Previous literature suggest that certain antidepressant medications are not as effective in patients with chronic systemic inflammation (Lanquillon, Krieg, Bening-Abu-Shach, & Vedder, 2000). For instance, a tumour necrosis factor antagonist (Raison et al. 2013) was tested as anti-depressant treatment and showed symptom improvement in patients with high baseline levels of CRP. Tumour necrosis factor also showed up in the DMP results of the area-based deprivation EWAS.

The EWAS results suggest that the underlying biological mechanisms of the cooccurrence of inflammatory disease and depression include exposure to environmental toxins, inflammatory mechanisms, liver functioning, kidney functioning, and glycogenesis. These findings could point towards aberrant HPA activity as the biological mechanism underlying the association between chronic systemic inflammation and depression.

5.3 Strengths and limitations

There are two major strengths of this thesis:

- 1) Providing an estimate of the risk of depression in low SES based on largest and most culturally inclusive systematic review and meta-analysis.
- 2) The use of novel and exploratory approaches to uncover the biological impact of multiple deprivation, inflammation and depression in the single largest dataset of its kind.

The DNAm dataset of Generation Scotland is the single largest of its kind and the analyses performed in this thesis is only the beginning. Even more DNAm data is being released at the time of writing and new advances are continually being made in DNAm based proteomic prediction. The Generation Scotland dataset is also deeply phenotyped and provides data linkage opportunities to pair up cohort data with routinely collected healthcare data. This incredible resource made it possible in the current work to correct for antidepressant medication use by linking prescription data to the existing DNAm and phenotypic data.

The systematic review and meta-analysis that lay the foundation for this thesis is the largest in terms of studies included and the most diverse in terms of countries represented. It has allowed almost universal conclusions about SES and depression risk and revealed that the magnitude of the risk is not necessarily dependent on the economic development of the country.

Another strength of the present study is the confirmation of results from previous literature as well as provision of new findings using novel and exploratory approaches. The present study supports previous studies showing evidence of an association between deprivation of accelerated epigenetic ageing and is the first to show this at the epigenome wide level. The attempt to explore the mechanisms underlying the link between inflammatory disease and depression, has added to the literature on a differential response to antidepressant medication in people with cooccurring depression and inflammatory disease.

This thesis has also provided possible support to the literature on childhood being biologically important for effects of deprivation; some of the differentially methylated genes in this thesis are preferentially expressed in early childhood and some of the gene enrichment pathways in this thesis are of an early developmental nature.

A translational strength of the present thesis has been to highlight possible targets for depression prevention and health improving interventions in general. Through meta-analyses and exploratory molecular approaches, this thesis has shown that income is a stronger risk factor than education or employment, as well as supporting results of previous research on early childhood being a particularly vulnerable time-period for detrimental effects of low SES. These findings together has the potential to guide a focused experimental interventions to further inform policy measures on preventing the development of depression. As mentioned in chapter 2 section 4.1 universal basic income has been trialled successfully in several countries as a tool to decrease equality and mitigate some of the harmful effects of living in low SES. The findings from this thesis, in particular the finding that income seems to be the strongest predictor for depression risk, lends support to the ideas of universal basic income and suggests that implementing such policies could have a positive effect on mental health at the population level.

One of the main limitations of the systematic review and meta-analysis is its high level of heterogeneity and the lack of a study quality measure. However, it is difficult to achieve a high level of diversity and a low level of heterogeneity at the same time. Assessing study quality is difficult in the case of observational studies because of the lack of a recognised tool for doing so. Future systematic reviews of observational or prevalence studies would greatly benefit from the development and recognition of such a tool. Another limitation is the potential for sex-specific effects, which the present thesis has not addressed. As mentioned in the general introduction, there are large sex differences in depression risk, and it is plausible that the effect of low SES on risk of depression is differs across sexes. The present thesis quantified the association between low SES and depression, future work would benefit the field by investigating any potential sex-specific effects.

A limitation of the two EWAS studies is the lack of ability to make conclusions on DNAm differences in brain tissue due to the use of whole blood DNAm data. Since this thesis focuses on the biological relationship between social environmental factors and a mental disorder it would have been a great addition to have DNAm from neural tissue.

However, there is a considerable overlap between methylation levels in brain tissue and in blood (Braun et al., 2019) and several of the enrichment analysis showed that differentially methylated genes were preferentially expressed in neuronal tissue or neuronal pathways. Furthermore, when dealing with chronic systemic inflammatory factors it may

make sense to look in blood and as mentioned earlier there is plenty of evidence that systemic chronic inflammation has effects on the brain as well.

Possible confounding from medications, particularly in the EWAS of area-based deprivation is also a limitation. Some of the stronger disease related signals from the DMPs and DMRs could come from medications changing levels DNAm. In one way this is not necessarily a problem because the analysis still picks up on biological differences across deprivation levels which was the aim of the study. However, the difference in strength could be due to some medications having a stronger impact on methylation levels. A differential impact may even be more specific; some medications could have a stronger impact on methylation levels in blood compared to other tissues. Finally, medication driven DNAm differences further limits any causal inferences, and it could even lead to bidirectional causality. The possibility of residual confounding from alcohol consumption, smoking and obesity, in the EWAS studies, is a limitation of this thesis. All three variables were included as covariates however some of the results unequivocally pointed to CpGs or genes previously associated with those three phenotypes. Residual confounding is likely to be a continuous presence in DNAm studies and will need careful methodological planning to avoid. Possible mitigations will be discussed in the future directions section. Another limitation of this thesis is the lack of external replication of the EWAS results. The prediction of deprivation in the ALSPAC cohort based on DNAm derived scores from Generation Scotland was not successful, despite a successful within cohort prediction from wave 1 into wave 3 of the Generation Scotland dataset. Similarly, the results from a previous EWAS into inflammatory disease and depression did not overlap with the results from this thesis. The lack of external replication is puzzling and the differences in DNAm patterns between cohort should be a focus of interrogation by future studies. It is possible that DNAm patterns are susceptible to geographical or population differences much the same way genetic markers are. It is also possible that small differences in the distribution of age or in the exact measurement of a phenotype from cohort-to-cohort impacts DNAm results much more than anticipated. These are all important points for future investigations.

When working with large cohort databases such as Generation Scotland it is important to be mindful of potential bias introduced because of sampling- and volunteer – processes. A now famous example is the UK Biobank, one of the most researched cohort datasets in the world, which has been found to generally be healthier, older wealthier and more likely to be female, than the general population (Fry et al. 2017). As mentioned in section Chapter 3 section 2 this is also, to some extent the case for Generation Scotland. These differences are

very likely due to volunteer bias, i.e. only 12.3% of contacted individuals volunteered to take part in the cohort. The fact that Generation Scotland is healthier than the general population is most likely even more apparent when it comes to comorbidities e.g. depression and inflammatory diseases. Individuals with comorbid diseases have a much higher disease burden and are therefore more likely to not be able to take part in a research study. This of course does not mean that the cohort is of no use to epidemiological research, however it is important to keep in mind when interpreting any findings from the cohort.

The Generation Scotland cohort was, from the start, designed as a family study. Family studies are immensely valuable for genetic epidemiology, investigating heritability, genetic versus environmental contributions to phenotypes as well as allowing for the study of parent-of-origin effects. However, the relatedness and cohabitation of individuals in a cohort can also present some limitations if not corrected for. For instance, shared living environments, and greater genetic overlap account for a substantial proportion of the variance in risk estimated for many disorders and will therefore introduce bias in analyses. This is also the case for the SES measure used in chapter 3 as it is an area-based measure. This limitation could however be turned to a strength as the data on family structures could be used to look into shared and non-shared environments as well as genetic contributions to differences in DNAm levels.

5.4 Future research

The present thesis has generated many possibilities for future research. First and foremost, it is pertinent to investigate whether chronic systemic inflammation mediates the relationship between low SES and increased risk of depression. To do this, using the Generation Scotland dataset, one could use DNAm derived CRP scores as an inflammatory mediating variable. Similar analyses could be done with other inflammatory markers such as IL-6 or GlycA. With new DNAm data being released there should be data of up to 18,000 individuals. A similar approach was recently employed by Conole et al. (2020) showing links between DNAm derived CRP score and brain ageing and Green et al. (2021) who found DNAm associated differences in brain regions previously implicated in depression.

Another interesting future direction would be to further interrogate the difference in association between subparts of SES and risk of depression. Based on the results of this thesis

I hypothesise that income is more strongly associated with depression risk than education or employment status. National registers in Nordic countries (Finland, Sweden, Norway and Denmark) provide great resources to test this association. They even allow for longitudinal studies, testing if a change in income makes a bigger difference in depression liability than a change in education or employment status.

Many of the phenotypes associated with differentially methylated genes or enriched pathways are either directly or indirectly implicated in the HPA axis and of course so is depression. Future research should specifically test whether HPA axis functioning mediates the relationship between low SES and increased risk of depression.

The lack of external but not internal replication in DNAm results warrants further investigation. Cross-sample robustness of DNAm patterns in general should be tested. And possible sources of any differences should be identified whether it is due to data processing, batch effects, geographical differences, or differences in lifestyle.

In general, epigenetic epidemiology is a relatively new area of research and an increased focus on what kind of questions we should and more importantly can ask with this kind of data is needed. Future studies should develop better way of covarying for phenotypes in DNAm studies. Residual confounding is likely to persist. To mitigate it we can use DNAm derived scores for phenotypes such as a smoking derived methylation score. However, until we design phenotypic data collection around the properties of DNAm such an approach may pick up on other sources of variation that could be valuable to study in some analyses.

For example, current vs. former smoker does not capture enough information to control for smoking induced differences in methylation. BMI is unlikely account for differences in obesity as it is too far removed from the biology of obesity that is picked up by DNAm data.

5.5 Conclusions

Higher prevalence of depression in low SES populations have been reported for centuries but large scale meta-analytical estimates of the difference in risk have been lacking or inconclusive. The biological mechanisms underlying the link between low SES and increased risk of depression, outwith known lifestyle factors, are poorly understood. The present thesis contributed with a cross-cultural estimate of depression risk associated with low SES and used epigenetic approaches to uncover the biological impact of living in low SES and possible biological mechanisms underlying the cooccurrence of depression and systemic chronic inflammation.

Results point to a robust increased risk of depression associated with low SES. Living in a deprived neighbourhood is associated with pervasive biological impacts of a primarily inflammatory nature. A combination of inflammatory diseases and depression is associated with early exposure to environmental toxins, differences in metabolic processes, liver and kidney functioning, depression treatment response, and educational and cognitive factors. Finally, gene set enrichment and pathway enrichment results indicate that early childhood may be a particularly vulnerable period for the biological impacts of deprivation and inflammation.

The findings of this thesis seem to point towards low SES being associated with increased risk of depression through a combination of increased psychological stress and systemic inflammation due to lifestyle-, and external environmental factors. These factors could cause a HPA imbalance which in turn could increase the risk of depression.

The results from the systematic review are congruent with the body of literature suggesting that relative poverty or deprivation is more important than absolute poverty.

Evidence from this thesis may be used to guide future research into preventing detrimental health effects of low SES. Results from this thesis indicate that efforts are best made at an early age, boosting income rather than employment and education, and lowering inequality may deliver the strongest results.

Bibliography

- Adams, N. L., Rose, T. C., Hawker, J., Violato, M., O'Brien, S. J., Barr, B., Howard, V. J. K., et al. (2018). Relationship between socioeconomic status and gastrointestinal infections in developed countries: A systematic review and meta-analysis. *Plos One*, 13(1), e0191633.
- Adkins, D. E., Khachane, A. N., McClay, J. L., Aberg, K., Bukszár, J., Sullivan, P. F., & van den Oord, E. J. C. G. (2012). SNP-based analysis of neuroactive ligand-receptor interaction pathways implicates PGE2 as a novel mediator of antipsychotic treatment response: data from the CATIE study. *Schizophrenia Research*, 135(1–3), 200–201.
- Adler, N. E., & Ostrove, J. M. (1999). Socioeconomic status and health: what we know and what we don't. *Annals of the New York Academy of Sciences*, 896, 3–15.
- Afeiche, M., Peterson, K. E., Sánchez, B. N., Schnaas, L., Cantonwine, D., Ettinger, A. S., Solano-González, M., et al. (2012). Windows of lead exposure sensitivity, attained height, and body mass index at 48 months. *The Journal of Pediatrics*, 160(6), 1044–1049.
- Agerbo, E., Sullivan, P. F., Vilhjálmsson, B. J., Pedersen, C. B., Mors, O., Børglum, A. D., Hougaard, D. M., et al. (2015). Polygenic Risk Score, Parental Socioeconomic Status, Family History of Psychiatric Disorders, and the Risk for Schizophrenia: A Danish Population-Based Study and Meta-analysis. *JAMA psychiatry*, 72(7), 635–641.
- Aizenstein, H. J., Baskys, A., Boldrini, M., Butters, M. A., Diniz, B. S., Jaiswal, M. K., Jellinger, K. A., et al. (2016). Vascular depression consensus report - a critical update. *BMC Medicine*, 14(1), 161.

- Al-Harbi, K. S. (2012). Treatment-resistant depression: therapeutic trends, challenges, and future directions. *Patient preference and adherence*, 6, 369–388.
- Alexopoulos, G. S., Meyers, B. S., Young, R. C., Campbell, S., Silbersweig, D., & Charlson, M. (1997). “Vascular depression” hypothesis. *Archives of General Psychiatry*, 54(10), 915–922.
- Ambugo, E. A. (2014). Cross-country variation in the sociodemographic factors associated with major depressive episode in Norway, the United Kingdom, Ghana, and Kenya. *Social Science & Medicine*, 113, 154–160.
- Andersen, A. M., Pietrzak, R. H., Kranzler, H. R., Ma, L., Zhou, H., Liu, X., Kramer, J., et al. (2017). Polygenic scores for major depressive disorder and risk of alcohol dependence. *JAMA psychiatry*, 74(11), 1153–1160.
- Anderson, J. O., Thundiyil, J. G., & Stolbach, A. (2012). Clearing the air: a review of the effects of particulate matter air pollution on human health. *Journal of Medical Toxicology*, 8(2), 166–175.
- Antczak, P., Ortega, F., Chipman, J. K., & Falciani, F. (2010). Mapping drug physico-chemical features to pathway activity reveals molecular networks linked to toxicity outcome. *Plos One*, 5(8), e12385.
- Antonsen, S., Mok, P. L. H., Webb, R. T., Mortensen, P. B., McGrath, J. J., Agerbo, E., Brandt, J., et al. (2020). Exposure to air pollution during childhood and risk of developing schizophrenia: a national cohort study. *The Lancet. Planetary Health*, 4(2), e64–e73.
- Arteaga-Henríquez, G., Simon, M. S., Burger, B., Weidinger, E., Wijkhuijs, A., Arolt, V., Birkenhager, T. K., et al. (2019). Low-Grade Inflammation as a Predictor of

- Antidepressant and Anti-Inflammatory Therapy Response in MDD Patients: A Systematic Review of the Literature in Combination With an Analysis of Experimental Data Collected in the EU-MOODINFLAME Consortium. *Frontiers in psychiatry*, 10, 458.
- Aryee, M. J., Jaffe, A. E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A. P., Hansen, K. D., & Irizarry, R. A. (2014). Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*, 30(10), 1363–1369.
- Austin, M. K., Chen, E., Ross, K. M., McEwen, L. M., Maclsaac, J. L., Kobor, M. S., & Miller, G. E. (2018). Early-life socioeconomic disadvantage, not current, predicts accelerated epigenetic aging of monocytes. *Psychoneuroendocrinology*, 97, 131–134.
- Awaworyi Churchill, S., & Farrell, L. (2017). Alcohol and depression: Evidence from the 2014 health survey for England. *Drug and Alcohol Dependence*, 180, 86–92.
- Bachmann, S. (2018). Epidemiology of suicide and the psychiatric perspective. *International Journal of Environmental Research and Public Health*, 15(7).
- Bahrami, S., Steen, N. E., Shadrin, A., O’Connell, K., Frei, O., Bettella, F., Wirgenes, K. V., et al. (2020). Shared Genetic Loci Between Body Mass Index and Major Psychiatric Disorders: A Genome-wide Association Study. *JAMA psychiatry*, 77(5), 503–512.
- Bai, P. Y., Wittert, G., Taylor, A. W., Martin, S. A., Milne, R. W., Jenkins, A. J., Januszewski, A. S., et al. (2017). The association between total phthalate concentration and non-communicable diseases and chronic inflammation in South Australian urban dwelling men. *Environmental Research*, 158, 366–372.

- Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: a practical tutorial. *Evidence-Based Mental Health*, 22(4), 153–160.
- Bao, C., Lv, J., Chen, J.-R., Wei, G.-Z., Liu, N., Wang, Y.-T., Ding, Z., et al. (2022). Chronic inflammation as a potential mediator between phthalate exposure and depressive symptoms. *Ecotoxicology and Environmental Safety*, 233, 113313.
- Barbu, Miruna C, Huider, F., Campbell, A., Amador, C., Adams, M. J., Lynall, M.-E., Howard, D. M., et al. (2021). Methylome-wide association study of antidepressant use in Generation Scotland and the Netherlands Twin Register implicates the innate immune system. *Molecular Psychiatry*.
- Barbu, Miruna C, Walker, R. M., Howard, D. M., Evans, K. L., Whalley, H. C., Porteous, D. J., Morris, S. W., et al. (2019). Epigenetic prediction of major depressive disorder. *medRxiv*.
- Barbu, M C, Campbell, A., Amador, C., Adams, M. J., Lynall, M. E., Howard, D. M., Walker, R. M., et al. (2020). Epigenome-wide association study of self-reported antidepressant use in Generation Scotland implicates the innate immune system. *medRxiv*.
- Barker, E. D., Walton, E., & Cecil, C. A. M. (2018). Annual Research Review: DNA methylation as a mediator in the association between risk exposure and child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 59(4), 303–322.
- Barnes, P. J., & Adcock, I. M. (2009). Glucocorticoid resistance in inflammatory diseases. *The Lancet*, 373(9678), 1905–1917.

- Battram, T., Yousefi, P., Crawford, G., Prince, C., Babei, M. S., Sharp, G., Hatcher, C., et al. (2021). The EWAS Catalog: a database of epigenome-wide association studies.
- Beard, E., Brown, J., West, R., Kaner, E., Meier, P., & Michie, S. (2019). Associations between socio-economic factors and alcohol consumption: A population survey of adults in England. *Plos One*, *14*(2), e0209442.
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*, 561–571.
- Beck, A. T. (1970). Cognitive therapy: Nature and relation to behavior therapy. *Behavior Therapy*, *1*(2), 184–200.
- Bellenguez, C., Küçükali, F., Jansen, I. E., Kleindam, L., Moreno-Grau, S., Amin, N., Naj, A. C., et al. (2022). New insights into the genetic etiology of Alzheimer’s disease and related dementias. *Nature Genetics*.
- Bengtsson, C., Nordmark, B., Klareskog, L., Lundberg, I., Alfredsson, L., & EIRA Study Group. (2005). Socioeconomic status and the risk of developing rheumatoid arthritis: results from the Swedish EIRA study. *Annals of the Rheumatic Diseases*, *64*(11), 1588–1594.
- Bennett, S., & Thomas, A. J. (2014). Depression and dementia: cause, consequence or coincidence? *Maturitas*, *79*(2), 184–190.
- Berk, M., Williams, L. J., Jacka, F. N., O’Neil, A., Pasco, J. A., Moylan, S., Allen, N. B., et al. (2013). So depression is an inflammatory disease, but where does the inflammation come from? *BMC Medicine*, *11*, 200.

- Bero, L., Chartres, N., Diong, J., Fabbri, A., Gherzi, D., Lam, J., Lau, A., et al. (2018). The risk of bias in observational studies of exposures (ROBINS-E) tool: concerns arising from application to observational studies of exposures. *Systematic Reviews*, 7(1), 242.
- Bhugra, D., & Ayonrinde, O. (2004). Depression in migrants and ethnic minorities. *Advances in Psychiatric Treatment*, 10(1), 13–17.
- Block, M. L., & Calderón-Garcidueñas, L. (2009). Air pollution: mechanisms of neuroinflammation and CNS disease. *Trends in Neurosciences*, 32(9), 506–516.
- Boden, J. M., & Fergusson, D. M. (2011). Alcohol and depression. *Addiction*, 106(5), 906–914.
- Border, R., Johnson, E. C., Evans, L. M., Smolen, A., Berley, N., Sullivan, P. F., & Keller, M. C. (2019). No Support for Historical Candidate Gene or Candidate Gene-by-Interaction Hypotheses for Major Depression Across Multiple Large Samples. *The American Journal of Psychiatry*, 176(5), 376–387.
- van den Bosch, M., & Meyer-Lindenberg, A. (2019). Environmental exposures and depression: biological mechanisms and epidemiological evidence. *Annual review of public health*, 40, 239–259.
- Bos, I., Jacobs, L., Nawrot, T. S., de Geus, B., Torfs, R., Int Panis, L., Degraeuwe, B., et al. (2011). No exercise-induced increase in serum BDNF after cycling near a major traffic road. *Neuroscience Letters*, 500(2), 129–132.
- Bouchard, M. F., Bellinger, D. C., Weuve, J., Matthews-Bellinger, J., Gilman, S. E., Wright, R. O., Schwartz, J., et al. (2009). Blood lead levels and major depressive disorder, panic

- disorder, and generalized anxiety disorder in US young adults. *Archives of General Psychiatry*, 66(12), 1313–1319.
- Boylan, J. M., & Robert, S. A. (2017). Neighborhood SES is particularly important to the cardiovascular health of low SES individuals. *Social Science & Medicine*, 188, 60–68.
- Braun, P. R., Han, S., Hing, B., Nagahama, Y., Gaul, L. N., Heinzman, J. T., Grossbach, A. J., et al. (2019). Genome-wide DNA methylation comparison between live human brain and peripheral tissues within individuals. *Translational psychiatry*, 9(1), 47.
- Braveman, P., & Gottlieb, L. (2014). The social determinants of health: it's time to consider the causes of the causes. *Public health reports (Washington, D.C. : 1974)*, 129 Suppl 2, 19–31.
- Brink, A. (1979). Depression and Loss: A Theme in Robert Burton's "Anatomy of Melancholy"(1621). *The Canadian Journal of Psychiatry*, 24(8), 767-772.
- Brook, R. D., Rajagopalan, S., Pope, C. A., Brook, J. R., Bhatnagar, A., Diez-Roux, A. V., Holguin, F., et al. (2010). Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation*, 121(21), 2331–2378.
- Bruce, M. L., Takeuchi, D. T., & Leaf, P. J. (1991). Poverty and psychiatric status. Longitudinal evidence from the New Haven Epidemiologic Catchment Area study. *Archives of General Psychiatry*, 48(5), 470–474.
- Bruijnzeel, A. W., & Gold, M. S. (2005). The role of corticotropin-releasing factor-like peptides in cannabis, nicotine, and alcohol dependence. *Brain Research. Brain Research Reviews*, 49(3), 505–528.

- Buniello, A., MacArthur, J. A. L., Cerezo, M., Harris, L. W., Hayhurst, J., Malangone, C., McMahon, A., et al. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Research*, 47(D1), D1005–D1012.
- Burke, H. M., Davis, M. C., Otte, C., & Mohr, D. C. (2005). Depression and cortisol responses to psychological stress: a meta-analysis. *Psychoneuroendocrinology*, 30(9), 846–856.
- Cai, N., Bigdeli, T. B., Kretschmar, W., Li, Y., Liang, J., Song, L., ... & Flint, J. (2015). Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature*, 523(7562), 588-591.
- Cai, S., Fan, J., Ye, J., Rao, X., & Li, Y. (2021). Phthalates exposure is associated with non-alcoholic fatty liver disease among US adults. *Ecotoxicology and Environmental Safety*, 224, 112665.
- Calderón-Garcidueñas, L., Calderón-Garcidueñas, A., Torres-Jardón, R., Avila-Ramírez, J., Kulesza, R. J., & Angiulli, A. D. (2015). Air pollution and your brain: what do you need to know right now. *Primary health care research & development*, 16(4), 329–345.
- Canela-Xandri, O., Law, A., Gray, A., Woolliams, J. A., & Tenesa, A. (2015). A new tool called DISSECT for analysing large genomic data sets using a Big Data approach. *Nature Communications*, 6, 10162.
- Cardenas, A., Rifas-Shiman, S. L., Agha, G., Hivert, M.-F., Litonjua, A. A., DeMeo, D. L., Lin, X., et al. (2017). Persistent DNA methylation changes associated with prenatal

mercury exposure and cognitive performance during childhood. *Scientific Reports*, 7(1), 288.

Carlsen, H. K., Forsberg, B., Meister, K., Gíslason, T., & Oudin, A. (2013). Ozone is associated with cardiopulmonary and stroke emergency hospital visits in Reykjavík, Iceland 2003-2009. *Environmental Health: A Global Access Science Source*, 12, 28.

Carroll, B. J., Feinberg, M., Greden, J. F., Tarika, J., Albala, A. A., Haskett, R. F., James, N. M., et al. (1981). A specific laboratory test for the diagnosis of melancholia. Standardization, validation, and clinical utility. *Archives of General Psychiatry*, 38(1), 15–22.

Cascade, E. F., Kalali, A. H., & Blier, P. (2007). Treatment of depression: antidepressant monotherapy and combination therapy. *Psychiatry (Edgmont)*, 4(11), 25–27.

Cereda, E., Sansone, V., Meola, G., & Malavazos, A. E. (2007). Increased visceral adipose tissue rather than BMI as a risk factor for dementia. *Age and Ageing*, 36(5), 488–491.

Chamberlain, S. R., Cavanagh, J., de Boer, P., Mondelli, V., Jones, D. N. C., Drevets, W. C., Cowen, P. J., et al. (2019). Treatment-resistant depression and peripheral C-reactive protein. *The British Journal of Psychiatry*, 214(1), 11–19.

Chandola, T., Brunner, E., & Marmot, M. (2006). Chronic stress at work and the metabolic syndrome: prospective study. *BMJ (Clinical Research Ed.)*, 332(7540), 521–525.

Chen, Y., Kassam, I., Lau, S. H., Kooner, J. S., Wilson, R., Peters, A., Winkelmann, J., et al. (2021). Impact of BMI and waist circumference on epigenome-wide DNA methylation and identification of epigenetic biomarkers in blood: an EWAS in multi-ethnic Asian individuals. *Clinical epigenetics*, 13(1), 195.

- Chuang, K.-J., Chan, C.-C., Su, T.-C., Lee, C.-T., & Tang, C.-S. (2007). The effect of urban air pollution on inflammation, oxidative stress, coagulation, and autonomic dysfunction in young adults. *American Journal of Respiratory and Critical Care Medicine*, 176(4), 370–376.
- Chuang, Y.-C., Cubbin, C., Ahn, D., & Winkleby, M. A. (2005). Effects of neighbourhood socioeconomic status and convenience store concentration on individual level smoking. *Journal of Epidemiology and Community Health*, 59(7), 568–573.
- Cipriani, A., Furukawa, T. A., Salanti, G., Chaimani, A., Atkinson, L. Z., Ogawa, Y., Leucht, S., et al. (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. *The Lancet*, 391(10128), 1357–1366.
- Cohen, S., Janicki-Deverts, D., Doyle, W. J., Miller, G. E., Frank, E., Rabin, B. S., & Turner, R. B. (2012). Chronic stress, glucocorticoid receptor resistance, inflammation, and disease risk. *Proceedings of the National Academy of Sciences of the United States of America*, 109(16), 5995–5999.
- Coker, E. S., Gunier, R., Huen, K., Holland, N., & Eskenazi, B. (2018). DNA methylation and socioeconomic status in a Mexican-American birth cohort. *Clinical epigenetics*, 10, 61.
- Colicino, E., Just, A., Kioumourtzoglou, M.-A., Vokonas, P., Cardenas, A., Sparrow, D., Weisskopf, M., et al. (2021). Blood DNA methylation biomarkers of cumulative lead exposure in adults. *Journal of Exposure Science & Environmental Epidemiology*, 31(1), 108–116.

- Collier, H. (1937). Health and poverty in Belfast. Report by the Ulster Society of Economic Research. *Ulster Medical Journal*.
- Cook. (2009). A Health Impact Assessment of the Healthy Families Act of 2009. *Oakland (CA): Human Impact Partners and San Francisco Department of Public Health*.
- Costa, L. G., Cole, T. B., Coburn, J., Chang, Y.-C., Dao, K., & Roqué, P. J. (2017). Neurotoxicity of traffic-related air pollution. *Neurotoxicology*, 59, 133–139.
- Crawford, B., Craig, Z., Mansell, G., White, I., Smith, A., Spaul, S., Imm, J., et al. (2018). DNA methylation and inflammation marker profiles associated with a history of depression. *Human Molecular Genetics*, 27(16), 2840–2850.
- Crum, R. M., Mojtabai, R., Lazareck, S., Bolton, J. M., Robinson, J., Sareen, J., Green, K. M., et al. (2013). A prospective assessment of reports of drinking to self-medicate mood symptoms with the incidence and persistence of alcohol dependence. *JAMA psychiatry*, 70(7), 718–726.
- Cuijpers, P., Vogelzangs, N., Twisk, J., Kleiboer, A., Li, J., & Penninx, B. W. (2014). Comprehensive meta-analysis of excess mortality in depression in the general community versus patients with specific illnesses. *The American Journal of Psychiatry*, 171(4), 453–462.
- Cuijpers, P. (2015). The combination of psychotherapy and pharmacotherapy in the treatment of.
- Cutler, D., Lleras-Muney, A., & Vogl, T. (2008). *Socioeconomic status and health: dimensions and mechanisms*. Cambridge, MA: National Bureau of Economic Research.

Dahl, E., & Birkelund, G. E. (1997). Health inequalities in later life in a social democratic welfare state. *Social Science & Medicine*, 44(6), 871–881.

Dazert, E., & Hall, M. N. (2011). mTOR signaling in disease. *Current Opinion in Cell Biology*, 23(6), 744–755.

Deierlein, A. L., Wolff, M. S., Pajak, A., Pinney, S. M., Windham, G. C., Galvez, M. P., Silva, M. J., et al. (2016). Longitudinal associations of phthalate exposures during childhood and body size measurements in young girls. *Epidemiology*, 27(4), 492–499.

De Matteis, S., Jarvis, D., Hutchings, S., Darnton, A., Fishwick, D., Sadhra, S., Rushton, L., et al. (2016). Occupations associated with COPD risk in the large population-based UK Biobank cohort study. *Occupational and Environmental Medicine*, 73(6), 378–384.

Depression and loss: a theme in Robert Burton’s “Anatomy of melancholy” (1621). - PDF Download Free. (n.d.). . Retrieved February 3, 2022, from https://docksci.com/depression-and-loss-a-theme-in-robert-burtons-anatomy-of-melancholy-1621_5daf9308097c4711638b4593.html

Depression, W. H. (2017). Other common mental disorders: global health estimates. *Geneva: World Health Organization*, 1–24.

DeRubeis, R J, Gelfand, L. A., Tang, T. Z., & Simons, A. D. (1999). Medications versus cognitive behavior therapy for severely depressed outpatients: mega-analysis of four randomized comparisons. *The American Journal of Psychiatry*, 156(7), 1007–1013.

DeRubeis, Robert J, Hollon, S. D., Amsterdam, J. D., Shelton, R. C., Young, P. R., Salomon, R. M., O’Reardon, J. P., et al. (2005). Cognitive therapy vs medications in the treatment of moderate to severe depression. *Archives of General Psychiatry*, 62(4), 409–416.

- Diez-Roux, A. V., Nieto, F. J., Muntaner, C., Tyroler, H. A., Comstock, G. W., Shahar, E., Cooper, L. S., et al. (1997). Neighborhood environments and coronary heart disease: A multilevel analysis. *American Journal of Epidemiology*, 146(1), 48–63.
- Dong, C., Sanchez, L. E., & Price, R. A. (2004). Relationship of obesity to depression: a family-based study. *International Journal of Obesity and Related Metabolic Disorders*, 28(6), 790–795.
- Dong, H., Zhang, W., Zeng, X., Hu, G., Zhang, H., He, S., & Zhang, S. (2014). Histamine induces upregulated expression of histamine receptors and increases release of inflammatory mediators from microglia. *Molecular Neurobiology*, 49(3), 1487–1500.
- Dunlop, B. W., LoParo, D., Kinkead, B., Mletzko-Crowe, T., Cole, S. P., Nemeroff, C. B., Mayberg, H. S., et al. (2019). Benefits of Sequentially Adding Cognitive-Behavioral Therapy or Antidepressant Medication for Adults With Nonremitting Depression. *The American Journal of Psychiatry*, 176(4), 275–286.
- Dunn, E. C., Brown, R. C., Dai, Y., Rosand, J., Nugent, N. R., Amstadter, A. B., & Smoller, J. W. (2015). Genetic determinants of depression: recent findings and future directions. *Harvard review of psychiatry*, 23(1), 1–18.
- Du, P., Kibbe, W. A., & Lin, S. M. (2008). lumi: a pipeline for processing Illumina microarray. *Bioinformatics*, 24(13), 1547–1548.
- Du, P., Zhang, X., Huang, C.-C., Jafari, N., Kibbe, W. A., Hou, L., & Lin, S. M. (2010). Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*, 11, 587.

- Ekman, M., Granström, O., Omérov, S., Jacob, J., & Landén, M. (2013). The societal cost of depression: evidence from 10,000 Swedish patients in psychiatric care. *Journal of Affective Disorders, 150*(3), 790–797.
- Enache, D., Pariante, C. M., & Mondelli, V. (2019). Markers of central inflammation in major depressive disorder: A systematic review and meta-analysis of studies examining cerebrospinal fluid, positron emission tomography and post-mortem brain tissue. *Brain, Behavior, and Immunity, 81*, 24–40.
- Enns, J. E., Nickel, N. C., Chartier, M., Chateau, D., Campbell, R., Phillips-Beck, W., Sarkar, J., et al. (2021). An unconditional prenatal income supplement is associated with improved birth and early childhood outcomes among First Nations children in Manitoba, Canada: a population-based cohort study. *BMC Pregnancy and Childbirth, 21*(1), 312.
- Evans, G. W., & Schamberg, M. A. (2009). Childhood poverty, chronic stress, and adult working memory. *Proceedings of the National Academy of Sciences of the United States of America, 106*(16), 6545–6549.
- Everson, S. A., Maty, S. C., Lynch, J. W., & Kaplan, G. A. (2002). Epidemiologic evidence for the relation between socioeconomic status and depression, obesity, and diabetes. *Journal of Psychosomatic Research, 53*(4), 891–895.
- First, M. (2016). User's guide for the SCID-5-CV Structured Clinical Interview for DSM-5® disorders: Clinical version. - PsycNET. Retrieved May 12, 2022, from <https://psycnet.apa.org/record/2016-15667-000>
- Flegal, K. M., Shepherd, J. A., Looker, A. C., Graubard, B. I., Borrud, L. G., Ogden, C. L., Harris, T. B., et al. (2009). Comparisons of percentage body fat, body mass index, waist

- circumference, and waist-stature ratio in adults. *The American Journal of Clinical Nutrition*, 89(2), 500–508.
- Fleiss JL. *The Design and Analysis of Clinical Experiments*. NY: Wiley and Sons; 1986
- Fluharty, M., Taylor, A. E., Grabski, M., & Munafò, M. R. (2017). The association of cigarette smoking with depression and anxiety: A systematic review. *Nicotine & Tobacco Research*, 19(1), 3–13.
- Fonken, L. K., Xu, X., Weil, Z. M., Chen, G., Sun, Q., Rajagopalan, S., & Nelson, R. J. (2011). Air pollution impairs cognition, provokes depressive-like behaviors and alters hippocampal cytokine expression and morphology. *Molecular Psychiatry*, 16(10), 987–95, 973.
- Forget, E. L. (2011). The Town with No Poverty: The Health Effects of a Canadian Guaranteed Annual Income Field Experiment. *Canadian Public Policy*, 37(3), 283–305.
- Fortin, J. P., Fertig, E., & Hansen, K. (2014). shinyMethyl: interactive quality control of Illumina 450k DNA methylation arrays in R [v2; ref status: indexed.
- Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., Heine-Suñer, D., et al. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences of the United States of America*, 102(30), 10604–10609.
- Franken, C., Lambrechts, N., Govarts, E., Koppen, G., Den Hond, E., Ooms, D., Voorspoels, S., et al. (2017). Phthalate-induced oxidative stress and association with asthma-related airway inflammation in adolescents. *International Journal of Hygiene and Environmental Health*, 220(2 Pt B), 468–477.

- Friedman, M. A., & Brownell, K. D. (1995). Psychological correlates of obesity: moving to the next research generation. *Psychological Bulletin*, 117(1), 3–20.
- Fried, E. I., & Nesse, R. M. (2015). Depression sum-scores don't add up: why analyzing specific depression symptoms is essential. *BMC Medicine*, 13, 72.
- Fry, A., Littlejohns, T. J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., ... & Allen, N. E. (2017). Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *American journal of epidemiology*, 186(9), 1026-1034.
- Furman, D., Campisi, J., Verdin, E., Carrera-Bastos, P., Targ, S., Franceschi, C., Ferrucci, L., et al. (2019). Chronic inflammation in the etiology of disease across the life span. *Nature Medicine*, 25(12), 1822–1832.
- Fu, P., & Yung, K. K. L. (2020). Air Pollution and Alzheimer's Disease: A Systematic Review and Meta-Analysis. *Journal of Alzheimer's Disease*, 77(2), 701–714.
- Gadd, D. A., Hillary, R. F., McCartney, D. L., Zaghlool, S. B., Stevenson, A. J., Cheng, Y., Fawns-Ritchie, C., et al. (2022). Epigenetic scores for the circulating proteome as tools for disease prediction. *eLife*, 11.
- Gardner, A., & Boles, R. G. (2011). Beyond the serotonin hypothesis: mitochondria, inflammation and neurodegeneration in major depression and affective spectrum disorders. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 35(3), 730–743.
- Gartlehner, G., Wagner, G., Matyas, N., Titscher, V., Greimel, J., Lux, L., Gaynes, B. N., et al. (2017). Pharmacological and non-pharmacological treatments for major depressive disorder: review of systematic reviews. *BMJ Open*, 7(6), e014912.

- GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, 392(10159), 1789–1858.
- Gilman, S. E., Kawachi, I., Fitzmaurice, G. M., & Buka, S. L. (2002). Socioeconomic status in childhood and the lifetime risk of major depression. *International Journal of Epidemiology*, 31(2), 359–367.
- Gilman, S. E., Sucha, E., Kingsbury, M., Horton, N. J., Murphy, J. M., & Colman, I. (2017). Depression and mortality in a longitudinal study: 1952-2011. *Canadian Medical Association Journal*, 189(42), E1304–E1310.
- Gold, S. M., Köhler-Forsberg, O., Moss-Morris, R., Mehnert, A., Miranda, J. J., Bullinger, M., Steptoe, A., et al. (2020). Comorbid depression in medical diseases. *Nature reviews. Disease primers*, 6(1), 69.
- Goldberg, D. (2011). The heterogeneity of “major depression”. *World psychiatry : official journal of the World Psychiatric Association (WPA)*, 10(3), 226–228.
- Goulden, R., Ibrahim, T., & Wolfson, C. (2015). Is high socioeconomic status a risk factor for multiple sclerosis? A systematic review. *European Journal of Neurology*, 22(6), 899–911.
- Greenberg, P. E., Fournier, A.-A., Sisitsky, T., Pike, C. T., & Kessler, R. C. (2015). The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *The Journal of Clinical Psychiatry*, 76(2), 155–162.

- Green, C., Shen, X., Stevenson, A. J., Conole, E. L. S., Harris, M. A., Barbu, M. C., Hawkins, E. L., et al. (2021). Structural brain correlates of serum and epigenetic markers of inflammation in major depressive disorder. *Brain, Behavior, and Immunity*, 92, 39–48.
- Green, M. J., & Benzeval, M. (2013). The development of socioeconomic inequalities in anxiety and depression symptoms over the lifecourse. *Social Psychiatry and Psychiatric Epidemiology*, 48(12), 1951–1961.
- Groffen, D. A. I., Koster, A., Bosma, H., van den Akker, M., Kempen, G. I. J. M., van Eijk, J. T. M., van Gool, C. H., et al. (2013). Unhealthy lifestyles do not mediate the relationship between socioeconomic status and incident depressive symptoms: the Health ABC study. *The American Journal of Geriatric Psychiatry*, 21(7), 664–674.
- Guo, B., Guo, Y., Nima, Q., Feng, Y., Wang, Z., Lu, R., Baimayangji, et al. (2022). Exposure to air pollution is associated with an increased risk of metabolic dysfunction-associated fatty liver disease. *Journal of Hepatology*, 76(3), 518–525.
- Guo, M., Mi, J., Jiang, Q.-M., Xu, J.-M., Tang, Y.-Y., Tian, G., & Wang, B. (2014). Metformin may produce antidepressant effects through improvement of cognitive function among depressed patients with diabetes mellitus. *Clinical and Experimental Pharmacology & Physiology*, 41(9), 650–656.
- Haas, H. L., Sergeeva, O. A., & Selbach, O. (2008). Histamine in the nervous system. *Physiological Reviews*, 88(3), 1183–1241.
- Hajat, A., Allison, M., Diez-Roux, A. V., Jenny, N. S., Jorgensen, N. W., Szpiro, A. A., Vedal, S., et al. (2015). Long-term exposure to air pollution and markers of inflammation,

- coagulation, and endothelial activation: a repeat-measures analysis in the Multi-Ethnic Study of Atherosclerosis (MESA). *Epidemiology*, 26(3), 310–320.
- Hakulinen, C., Musliner, K. L., & Agerbo, E. (2019). Bipolar disorder and depression in early adulthood and long-term employment, income, and educational attainment: A nationwide cohort study of 2,390,127 individuals. *Depression and Anxiety*, 36(11), 1080–1088.
- Hammen, C, Davila, J., Brown, G., Ellicott, A., & Gitlin, M. (1992). Psychiatric history and stress: predictors of severity of unipolar depression. *Journal of Abnormal Psychology*, 101(1), 45–52.
- Hammen, Constance. (2005). Stress and depression. *Annual Review of Clinical Psychology*, 1, 293–319.
- Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., Sadda, S., Klotzle, B., et al. (2013). Genome-wide methylation profiles reveal quantitative views of human aging rates. *Molecular Cell*, 49(2), 359–367.
- Han, C., Lim, Y.-H., & Hong, Y.-C. (2016). Does cadmium exposure contribute to depressive symptoms in the elderly population? *Occupational and Environmental Medicine*, 73(4), 269–274.
- Han, L. K. M., Aghajani, M., Clark, S. L., Chan, R. F., Hattab, M. W., Shabalin, A. A., Zhao, M., et al. (2018). Epigenetic aging in major depressive disorder. *The American Journal of Psychiatry*, 175(8), 774–782.
- Harding, K. E., Wardle, M., Carruthers, R., Robertson, N., Zhu, F., Kingwell, E., & Tremlett, H. (2019). Socioeconomic status and disability progression in multiple sclerosis: A multinational study. *Neurology*, 92(13), e1497–e1506.

- Harmer, C. J., Duman, R. S., & Cowen, P. J. (2017). How do antidepressants work? New perspectives for refining future treatment approaches. *The Lancet. Psychiatry*, 4(5), 409–418.
- Harper, K. N., Peters, B. A., & Gamble, M. V. (2013). Batch Effects and Pathway Analysis: Two Potential Perils in Cancer Studies Involving DNA Methylation Array Analysis Batch Effects, Pathway Analysis, and DNA Methylation Arrays. *Cancer epidemiology, biomarkers & prevention*, 22(6), 1052-1060.
- Harrison, S., Davies, A. R., Dickson, M., Tyrrell, J., Green, M. J., Katikireddi, S. V., Campbell, D., et al. (2020). The causal effects of health conditions and risk factors on social and socioeconomic outcomes: Mendelian randomization in UK Biobank. *International Journal of Epidemiology*, 49(5), 1661–1681.
- Heard, E., & Martienssen, R. A. (2014). Transgenerational epigenetic inheritance: myths and mechanisms. *Cell*, 157(1), 95–109.
- Heils, A., Teufel, A., Petri, S., Stöber, G., Riederer, P., Bengel, D., & Lesch, K. P. (1996). Allelic variation of human serotonin transporter gene expression. *Journal of Neurochemistry*, 66(6), 2621–2624.
- Hersey, M., Hashemi, P., & Reagan, L. P. (2021). Integrating the monoamine and cytokine hypotheses of depression: Is histamine the missing link? *The European Journal of Neuroscience*.
- Hertzman, C. (1999). The biological embedding of early experience and its effects on health in adulthood. *Annals of the New York Academy of Sciences*, 896, 85–95.

- Hillary, R. F., Stevenson, A. J., Cox, S. R., McCartney, D. L., Harris, S. E., Seeboth, A., Higham, J., et al. (2021). An epigenetic predictor of death captures multi-modal measures of brain health. *Molecular Psychiatry*, 26(8), 3806–3816.
- Hiscock, R., Bauld, L., Amos, A., Fidler, J. A., & Munafò, M. (2012). Socioeconomic status and smoking: a review. *Annals of the New York Academy of Sciences*, 1248, 107–123.
- Hlisníková, H., Petrovičová, I., Kolena, B., Šidlovská, M., & Sirotkin, A. (2021). Effects and mechanisms of phthalates' action on neurological processes and neural health: a literature review. *Pharmacological Reports*, 73(2), 386–404.
- Hopcroft, R. L., & Bradley, D. B. (2007). The sex difference in depression across 29 countries. *Social Forces*, 85(4), 1483–1507.
- Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome Biology*, 14(10), R115.
- Hosseinpoor, A. R., Bergen, N., Mendis, S., Harper, S., Verdes, E., Kunst, A., & Chatterji, S. (2012). Socioeconomic inequality in the prevalence of noncommunicable diseases in low- and middle-income countries: results from the World Health Survey. *BMC Public Health*, 12, 474.
- Howard, D. M., Adams, M. J., Clarke, T.-K., Hafferty, J. D., Gibson, J., Shirali, M., Coleman, J. R. I., et al. (2019). Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature Neuroscience*, 22(3), 343–352.

- Humphreys, K. L., LeMoult, J., Wear, J. G., Piersiak, H. A., Lee, A., & Gotlib, I. H. (2020). Child maltreatment and depression: A meta-analysis of studies using the Childhood Trauma Questionnaire. *Child Abuse & Neglect*, 102, 104361.
- Husain, M. I., Strawbridge, R., Stokes, P. R., & Young, A. H. (2017). Anti-inflammatory treatments for mood disorders: Systematic review and meta-analysis. *Journal of Psychopharmacology*, 31(9), 1137–1148.
- Hu, X. H., Bull, S. A., Hunkeler, E. M., Ming, E., Lee, J. Y., Fireman, B., & Markson, L. E. (2004). Incidence and duration of side effects and those rated as bothersome with selective serotonin reuptake inhibitor treatment for depression: patient report versus physician estimate. *The Journal of Clinical Psychiatry*, 65(7), 959–965.
- Hyde, C. L., Nagle, M. W., Tian, C., Chen, X., Paciga, S. A., Wendland, J. R., ... & Winslow, A. R. (2016). Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nature genetics*, 48(9), 1031-1036.
- IntHout, J., Ioannidis, J. P. A., & Borm, G. F. (2014). The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Medical Research Methodology*, 14, 25.
- Jacobson, L., & Sapolsky, R. (1991). The role of the hippocampus in feedback regulation of the hypothalamic-pituitary-adrenocortical axis. *Endocrine Reviews*, 12(2), 118–134.
- Jacobson, M. H., Stein, C. R., Liu, M., Ackerman, M. G., Blakemore, J. K., Long, S. E., Pinna, G., et al. (2021). Prenatal exposure to bisphenols and phthalates and postpartum depression: the role of neurosteroid hormone disruption. *The Journal of Clinical Endocrinology and Metabolism*, 106(7), 1887–1899.

- Jaffe, D. H., Rive, B., & Dennee, T. R. (2019). The humanistic and economic burden of treatment-resistant depression in Europe: a cross-sectional study. *BMC Psychiatry*, 19(1), 247.
- James-Todd, T., Tehranifar, P., Rich-Edwards, J., Titievsky, L., & Terry, M. B. (2010). The impact of socioeconomic status across early life on age at menarche among a racially diverse population of girls. *Annals of Epidemiology*, 20(11), 836–842.
- Jarman, B., Townsend, P., & Carstairs, V. (1991). Deprivation indices. *BMJ (Clinical Research Ed.)*, 303(6801), 523.
- Ji, Y., Schaid, D. J., Desta, Z., Kubo, M., Batzler, A. J., Snyder, K., Mushiroda, T., et al. (2014). Citalopram and escitalopram plasma drug and metabolite concentrations: genome-wide associations. *British Journal of Clinical Pharmacology*, 78(2), 373–383.
- Jobst, A., Brakemeier, E. L., Buchheim, A., Caspar, F., Cuijpers, P., Ebmeier, K. P., Falkai, P., et al. (2016). European Psychiatric Association Guidance on psychotherapy in chronic depression across Europe. *European Psychiatry*, 33, 18–36.
- Joehanes, R., Just, A. C., Marioni, R. E., Pilling, L. C., Reynolds, L. M., Mandaviya, P. R., Guan, W., et al. (2016). Epigenetic signatures of cigarette smoking. *Circulation. Cardiovascular Genetics*, 9(5), 436–447.
- Juruena, M. F. (2014). Early-life stress and HPA axis trigger recurrent adulthood depression. *Epilepsy & Behavior*, 38, 148–159.
- Kappelmann, N., Lewis, G., Dantzer, R., Jones, P. B., & Khandaker, G. M. (2018). Antidepressant activity of anti-cytokine treatment: a systematic review and meta-analysis

of clinical trials of chronic inflammatory conditions. *Molecular Psychiatry*, 23(2), 335–343.

Kasper Daniel Hansen [Cre, A. (2017). IlluminaHumanMethylationEPICanno.ilm10b2.hg19. *Bioconductor*.

Kaysen, G. A. (2009). Biochemistry and biomarkers of inflamed patients: why look, what to assess. *Clinical Journal of the American Society of Nephrology*, 4 Suppl 1, S56-63.

Kelly, K., Posternak, M., & Alpert, J. E. (2008). Toward achieving optimal response: understanding and managing antidepressant side effects. *Dialogues in Clinical Neuroscience*, 10(4), 409–418.

Kern, J. K., Geier, D. A., Bjørklund, G., King, P. G., Homme, K. G., Haley, B. E., Sykes, L. K., et al. (2014). Evidence supporting a link between dental amalgams and chronic illness, fatigue, depression, anxiety, and suicide. *Neuro Endocrinology Letters*, 35(7), 537–552.

Kessler, R. C., & Bromet, E. J. (2013). The epidemiology of depression across cultures. *Annual review of public health*, 34, 119–138.

Kessler, R. C. (2012). The costs of depression. *The Psychiatric Clinics of North America*, 35(1), 1–14.

Khandaker, G. M., Pearson, R. M., Zammit, S., Lewis, G., & Jones, P. B. (2014). Association of serum interleukin 6 and C-reactive protein in childhood with depression and psychosis in young adult life: a population-based longitudinal study. *JAMA psychiatry*, 71(10), 1121–1128.

- Kim, H.-J., Min, J.-Y., Seo, Y.-S., & Min, K.-B. (2019). Association of Ambient Air Pollution with Increased Liver Enzymes in Korean Adults. *International Journal of Environmental Research and Public Health*, 16(7).
- Köhler, C. A., Evangelou, E., Stubbs, B., Solmi, M., Veronese, N., Belbasis, L., Bortolato, B., et al. (2018). Mapping risk factors for depression across the lifespan: An umbrella review of evidence from meta-analyses and Mendelian randomization studies. *Journal of Psychiatric Research*, 103, 189–207.
- Koistinen, P., & Perkiö, J. (2014). Good and bad times of social innovations: the case of universal basic income in finland. *Basic Income Studies*, 9(1–2).
- Kozyrskyj, A. L., Kendall, G. E., Jacoby, P., Sly, P. D., & Zubrick, S. R. (2010). Association between socioeconomic status and the development of asthma: analyses of income trajectories. *American Journal of Public Health*, 100(3), 540–546.
- Kuźma, Ł., Małyszko, J., Bachórzewska-Gajewska, H., Kralisz, P., & Dobrzycki, S. (2021). Exposure to air pollution and renal function. *Scientific Reports*, 11(1), 11419.
- Kyrou, I., Tsigos, C., Mavrogianni, C., Cardon, G., Van Stappen, V., Latomme, J., Kivelä, J., et al. (2020). Sociodemographic and lifestyle-related risk factors for identifying vulnerable groups for type 2 diabetes: a narrative review with emphasis on data from Europe. *BMC endocrine disorders*, 20(Suppl 1), 134.
- Lanquillon, S., Krieg, J. C., Bening-Abu-Shach, U., & Vedder, H. (2000). Cytokine production and treatment response in major depressive disorder. *Neuropsychopharmacology*, 22(4), 370–379.

- Lee, J. W., Ko, J., Ju, C., & Eltzschig, H. K. (2019). Hypoxia signaling in human diseases and therapeutic targets. *Experimental & Molecular Medicine*, 51(6), 1–13.
- Lee, K.-S., Lim, Y.-H., Kim, K.-N., Choi, Y.-H., Hong, Y.-C., & Lee, N. (2018). Urinary phthalate metabolites concentrations and symptoms of depression in an elderly population. *The Science of the Total Environment*, 625, 1191–1197.
- Li, M., D'Arcy, C., & Meng, X. (2016). Maltreatment in childhood substantially increases the risk of adult depression and anxiety in prospective cohort studies: systematic review, meta-analysis, and proportional attributable fractions. *Psychological Medicine*, 46(4), 717–730.
- Lindqvist, D., Dhabhar, F. S., James, S. J., Hough, C. M., Jain, F. A., Bersani, F. S., Reus, V. I., et al. (2017). Oxidative stress, inflammation and treatment response in major depression. *Psychoneuroendocrinology*, 76, 197–205.
- Li, Y., Xu, L., Shan, Z., Teng, W., & Han, C. (2019). Association between air pollution and type 2 diabetes: an updated review of the literature. *Therapeutic advances in endocrinology and metabolism*, 10, 2042018819897046.
- Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P., & Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Systems*, 1(6), 417–425.
- Lidsky, T. I., & Schneider, J. S. (2003). Lead neurotoxicity in children: basic mechanisms and clinical correlates. *Brain: A Journal of Neurology*, 126(Pt 1), 5–19.
- Ligthart, S., Marzi, C., Aslibekyan, S., Mendelson, M. M., Conneely, K. N., Tanaka, T., Colicino, E., et al. (2016). DNA methylation signatures of chronic low-grade inflammation are associated with complex diseases. *Genome Biology*, 17(1), 255.

- Lim, Y.-H., Kim, H., Kim, J. H., Bae, S., Park, H. Y., & Hong, Y.-C. (2012). Air pollution and symptoms of depression in elderly adults. *Environmental Health Perspectives*, 120(7), 1023–1028.
- Lindqvist, D., Janelidze, S., Hagell, P., Erhardt, S., Samuelsson, M., Minthon, L., Hansson, O., et al. (2009). Interleukin-6 is elevated in the cerebrospinal fluid of suicide attempters and related to symptom severity. *Biological Psychiatry*, 66(3), 287–292.
- Liu, C., Marioni, R. E., Hedman, Å. K., Pfeiffer, L., Tsai, P. C., Reynolds, L. M., Just, A. C., et al. (2018). A DNA methylation biomarker of alcohol consumption. *Molecular Psychiatry*, 23(2), 422–433.
- Lleras-Muney, A. (2005). The relationship between education and adult mortality in the United States. *The Review of economic studies*, 72(1), 189–221.
- Lofors, J., & Sundquist, K. (2007). Low-linking social capital as a predictor of mental disorders: a cohort study of 4.5 million Swedes. *Social Science & Medicine*, 64(1), 21–34.
- Lohoff, F. W., Clarke, T.-K., Kaminsky, Z. A., Walker, R. M., Bermingham, M. L., Jung, J., Morris, S. W., et al. (2021). Epigenome-wide association study of alcohol consumption in N = 8161 individuals and relevance to alcohol use disorder pathophysiology: identification of the cystine/glutamate transporter SLC7A11 as a top target. *Molecular Psychiatry*.
- Long, M., Zhang, C., Xu, D., Fu, W., Gan, X., Li, F., Wang, Q., et al. (2019). PM2.5 aggravates diabetes via the systemically activated IL-6-mediated STAT3/SOCS3 pathway in rats' liver. *Environmental Pollution*, 113342.

- Lorant, V., Delière, D., Eaton, W., Robert, A., Philippot, P., & Ansseau, M. (2003). Socioeconomic inequalities in depression: a meta-analysis. *American Journal of Epidemiology*, 157(2), 98–112.
- Lund, C., Breen, A., Flisher, A. J., Kakuma, R., Corrigall, J., Joska, J. A., Swartz, L., et al. (2010). Poverty and common mental disorders in low and middle income countries: A systematic review. *Social Science & Medicine*, 71(3), 517–528.
- Lupien, Sonia J, McEwen, B. S., Gunnar, M. R., & Heim, C. (2009). Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nature Reviews. Neuroscience*, 10(6), 434–445.
- Lupien, S J, Maheu, F., Tu, M., Fiocco, A., & Schramek, T. E. (2007). The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. *Brain and Cognition*, 65(3), 209–237.
- Lu, A. T., Quach, A., Wilson, J. G., Reiner, A. P., Aviv, A., Raj, K., Hou, L., et al. (2019). DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging*, 11(2), 303–327.
- Lynch, J. W., Kaplan, G. A., & Salonen, J. T. (1997). Why do poor people behave poorly? Variation in adult health behaviours and psychosocial characteristics by stages of the socioeconomic lifecourse. *Social Science & Medicine*, 44(6), 809–819.
- Lyra E Silva, N. de M., Lam, M. P., Soares, C. N., Munoz, D. P., Milev, R., & De Felice, F. G. (2019). Insulin resistance as a shared pathogenic mechanism between depression and type 2 diabetes. *Frontiers in psychiatry*, 10, 57.

- Lyst, M. J., & Bird, A. (2015). Rett syndrome: a complex disorder with simple roots. *Nature Reviews. Genetics*, 16(5), 261–275.
- Macintyre, S., McKay, L., Cummins, S., & Burns, C. (2005). Out-of-home food outlets and area deprivation: case study in Glasgow, UK. *The International Journal of Behavioral Nutrition and Physical Activity*, 2, 16.
- Madden, R. A., McCartney, D. L., Walker, R. M., Hillary, R. F., Bermingham, M. L., Rawlik, K., Morris, S. W., et al. (2021). Birth weight associations with DNA methylation differences in an adult population. *Epigenetics*, 16(7), 783–796.
- Malgaroli, M., Calderon, A., & Bonanno, G. A. (2021). Networks of major depressive disorder: A systematic review. *Clinical Psychology Review*, 85, 102000.
- Marini, S., Davis, K. A., Soare, T. W., Zhu, Y., Suderman, M. J., Simpkin, A. J., Smith, A. D. A. C., et al. (2020). Adversity exposure during sensitive periods predicts accelerated epigenetic aging in children. *Psychoneuroendocrinology*, 113, 104484.
- Marmot. (2020). Health Equity in England: The Marmot Review 10 Years On | The Health Foundation. Retrieved April 30, 2020, from <https://www.health.org.uk/publications/reports/the-marmot-review-10-years-on>
- Marmot, M. G., Shipley, M. J., & Rose, G. (1984). Inequalities in death--specific explanations of a general pattern? *The Lancet*, 1(8384), 1003–1006.
- Marmot, M. G., Smith, G. D., Stansfeld, S., Patel, C., North, F., Head, J., White, I., et al. (1991). Health inequalities among British civil servants: the Whitehall II study. *The Lancet*, 337(8754), 1387–1393.

- Marttila, S., Kananen, L., Häyrynen, S., Jylhävä, J., Nevalainen, T., Hervonen, A., Jylhä, M., et al. (2015). Ageing-associated changes in the human DNA methylome: genomic locations and effects on gene expression. *BMC Genomics*, *16*, 179.
- McCartney, D. L., Stevenson, A. J., Hillary, R. F., Walker, R. M., Bermingham, M. L., Morris, S. W., Clarke, T.-K., et al. (2018). Epigenetic signatures of starting and stopping smoking. *EBioMedicine*, *37*, 214–220.
- McCartney, D. L., Walker, R. M., Morris, S. W., McIntosh, A. M., Porteous, D. J., & Evans, K. L. (2016). Identification of polymorphic and off-target probe binding sites on the Illumina Infinium MethylationEPIC BeadChip. *Genomics data*, *9*, 22–24.
- McDade, T. W., Ryan, C. P., Jones, M. J., Hoke, M. K., Borja, J., Miller, G. E., Kuzawa, C. W., et al. (2019). Genome-wide analysis of DNA methylation in relation to socioeconomic status during development and early adulthood. *American Journal of Physical Anthropology*, *169*(1), 3–11.
- McDermott, E., Ryan, E. J., Tosetto, M., Gibson, D., Burrage, J., Keegan, D., Byrne, K., et al. (2016). DNA Methylation Profiling in Inflammatory Bowel Disease Provides New Insights into Disease Pathogenesis. *Journal of Crohn's & colitis*, *10*(1), 77–86.
- McEwen, B. S., & Gianaros, P. J. (2010). Central role of the brain in stress and adaptation: links to socioeconomic status, health, and disease. *Annals of the New York Academy of Sciences*, *1186*, 190–222.
- McGuinness, D., McGlynn, L. M., Johnson, P. C. D., MacIntyre, A., Batty, G. D., Burns, H., Cavanagh, J., et al. (2012). Socio-economic status is associated with epigenetic differences in the pSoBid cohort. *International Journal of Epidemiology*, *41*(1), 151–160.

- McInnes, L. A., & Freimer, N. B. (1995). Mapping genes for psychiatric disorders and behavioral traits. *Current Opinion in Genetics & Development*, 5(3), 376–381.
- McLaren, L. (2007). Socioeconomic status and obesity. *Epidemiologic Reviews*, 29, 29–48.
- McManus, S., Meltzer, H., & Campion, J. (2009). Cigarette smoking and mental health in England: Data from the Adult Psychiatric Morbidity Survey 2007.
- Menkin, V. (1943). Gluconeogenesis and cellular injury. a further inquiry into the mechanism involved in diabetes enhanced by inflammation. *American Journal of Physiology-Legacy Content*, 138(3), 396–407.
- de Mestral, C., & Stringhini, S. (2017). Socioeconomic status and cardiovascular disease: an update. *Current cardiology reports*, 19(11), 115.
- Milaneschi, Y., Simmons, W. K., van Rossum, E. F. C., & Penninx, B. W. (2019). Depression and obesity: evidence of shared biological mechanisms. *Molecular Psychiatry*, 24(1), 18–33.
- Milaniak, I., & Jaffee, S. R. (2019). Childhood socioeconomic status and inflammation: A systematic review and meta-analysis. *Brain, Behavior, and Immunity*, 78, 161–176.
- Miller, A. H., & Raison, C. L. (2016). The role of inflammation in depression: from evolutionary imperative to modern treatment target. *Nature Reviews. Immunology*, 16(1), 22–34.
- Misiak, B., Stańczykiewicz, B., Pawlak, A., Szewczuk-Bogusławska, M., Samochowiec, J., Samochowiec, A., Tyburski, E., et al. (2022). Adverse childhood experiences and low

- socioeconomic status with respect to allostatic load in adulthood: A systematic review. *Psychoneuroendocrinology*, 136, 105602.
- Miszkiewicz, J. J., & Cooke, K. M. (2019). Socio-economic Determinants of Bone Health from Past to Present. *Clinical reviews in bone and mineral metabolism*, 17(3–4), 109–122.
- Mitchell, A. J., Chan, M., Bhatti, H., Halton, M., Grassi, L., Johansen, C., & Meader, N. (2011). Prevalence of depression, anxiety, and adjustment disorder in oncological, haematological, and palliative-care settings: a meta-analysis of 94 interview-based studies. *The Lancet Oncology*, 12(2), 160–174.
- Mrazek, D. A., Hornberger, J. C., Altar, C. A., & Degtiar, I. (2014). A review of the clinical, economic, and societal burden of treatment-resistant depression: 1996-2013. *Psychiatric Services*, 65(8), 977–987.
- Mulder, R. H., Neumann, A., Cecil, C. A. M., Walton, E., Houtepen, L. C., Simpkin, A. J., Rijlaarsdam, J., et al. (2021). Epigenome-wide change and variation in DNA methylation in childhood: trajectories from birth to late adolescence. *Human Molecular Genetics*, 30(1), 119–134.
- Muslimovic, A., Tulumovic, D., Hasanspahic, S., Hamzic-Mehmedbasic, A., & Temimovi, R. (2015). Serum cystatin C - marker of inflammation and cardiovascular morbidity in chronic kidney disease stages 1-4. *Materia socio-medica*, 27(2), 75–78.
- Nandi, A., Glymour, M. M., & Subramanian, S. V. (2014). Association among socioeconomic status, health behaviors, and all-cause mortality in the United States. *Epidemiology*, 25(2), 170–177.

- Navrady, L. B., Wolters, M. K., MacIntyre, D. J., Clarke, T. K., Campbell, A. I., Murray, A. D., Evans, K. L., et al. (2018). Cohort Profile: Stratifying Resilience and Depression Longitudinally (STRADL): a questionnaire follow-up of Generation Scotland: Scottish Family Health Study (GS:SFHS). *International Journal of Epidemiology*, 47(1), 13–14g.
- Newell-Price, J., Clark, A. J., & King, P. (2000). DNA methylation and silencing of gene expression. *Trends in Endocrinology & Metabolism*, 11(4), 142-148.
- Ng, J.-S., & Chin, K.-Y. (2021). Potential mechanisms linking psychological stress to bone health. *International Journal of Medical Sciences*, 18(3), 604–614.
- Nuttall, F. Q. (2015). Body mass index: obesity, BMI, and health: A critical review. *Nutrition today*, 50(3), 117–128.
- Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S. M., Sidorenko, J., et al. (2022). Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nature Genetics*, 54(4), 437–449.
- Okin, D., & Medzhitov, R. (2012). Evolution of inflammatory diseases. *Current Biology*, 22(17), R733-40.
- Palinkas, L. A., Wingard, D. L., & Barrett-Connor, E. (1996). Depressive symptoms in overweight and obese older adults: a test of the “jolly fat” hypothesis. *Journal of Psychosomatic Research*, 40(1), 59–66.
- Permanyer, I., & Smits, J. (2020). Inequality in Human Development across the Globe. *Population and development review*, 46(3), 583–601.

- Pidsley, R., Y Wong, C. C., Volta, M., Lunnon, K., Mill, J., & Schalkwyk, L. C. (2013). A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics*, 14, 293.
- Plana-Ripoll, O., Pedersen, C. B., Holtz, Y., Benros, M. E., Dalsgaard, S., de Jonge, P., Fan, C. C., et al. (2019). Exploring comorbidity within mental disorders among a danish national population. *JAMA psychiatry*, 76(3), 259–270.
- Polderman, T. J. C., Benyamin, B., de Leeuw, C. A., Sullivan, P. F., van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, 47(7), 702–709.
- Pollack, C. E., Cubbin, C., Ahn, D., & Winkleby, M. (2005). Neighbourhood deprivation and alcohol consumption: does the availability of alcohol play a role? *International Journal of Epidemiology*, 34(4), 772–780.
- Prescott, E., & Vestbo, J. (1999). Socioeconomic status and chronic obstructive pulmonary disease. *Thorax*, 54(8), 737–741.
- Ragguett, R.-M., Cha, D. S., Subramaniapillai, M., Carmona, N. E., Lee, Y., Yuan, D., Rong, C., et al. (2017). Air pollution, aeroallergens and suicidality: a review of the effects of air pollution and aeroallergens on suicidal behavior and an exploration of possible mechanisms. *Reviews on environmental health*, 32(4), 343–359.
- Raison, Charles L, Rutherford, R. E., Woolwine, B. J., Shuo, C., Schettler, P., Drake, D. F., Haroon, E., et al. (2013). A randomized controlled trial of the tumor necrosis factor antagonist infliximab for treatment-resistant depression: the role of baseline inflammatory biomarkers. *JAMA psychiatry*, 70(1), 31–41.

- Raison, C L, & Miller, A. H. (2013). The evolutionary significance of depression in Pathogen Host Defense (PATHOS-D). *Molecular Psychiatry*, 18(1), 15–37.
- Ramachandrai, C. T., Subramanyam, N., Bar, K. J., Baker, G., & Yeragani, V. K. (2011). Antidepressants: From MAOIs to SSRIs and more. *Indian journal of psychiatry*, 53(2), 180–182.
- Reidpath, D. D., Burns, C., Garrard, J., Mahoney, M., & Townsend, M. (2002). An ecological study of the relationship between social and environmental determinants of obesity. *Health & Place*, 8(2), 141–145.
- Reiss, F. (2013). Socioeconomic inequalities and mental health problems in children and adolescents: a systematic review. *Social Science & Medicine*, 90, 24–31.
- Relton, C. L., Gaunt, T., McArdle, W., Ho, K., Duggirala, A., Shihab, H., Woodward, G., et al. (2015). Data resource profile: accessible resource for integrated epigenomic studies (ARIES). *International Journal of Epidemiology*, 44(4), 1181–1190.
- Richardson, R., Westley, T., Gariépy, G., Austin, N., & Nandi, A. (2015). Neighborhood socioeconomic conditions and depression: a systematic review and meta-analysis. *Social Psychiatry and Psychiatric Epidemiology*, 50(11), 1641–1656.
- Robbins, J. M., Vaccarino, V., Zhang, H., & Kasl, S. V. (2005). Socioeconomic status and diagnosed diabetes incidence. *Diabetes Research and Clinical Practice*, 68(3), 230–236.
- Roberts, Robert E, Strawbridge, W. J., Deleger, S., & Kaplan, G. A. (2002). Are the fat more jolly? *Annals of Behavioral Medicine*, 24(3), 169–180.

- Roberts, R E, Kaplan, G. A., Shema, S. J., & Strawbridge, W. J. (2000). Are the obese at greater risk for depression? *American Journal of Epidemiology*, 152(2), 163–170.
- Rosengren, A., Smyth, A., Rangarajan, S., Ramasundarahettige, C., Bangdiwala, S. I., AlHabib, K. F., Avezum, A., et al. (2019). Socioeconomic status and risk of cardiovascular disease in 20 low-income, middle-income, and high-income countries: the Prospective Urban Rural Epidemiologic (PURE) study. *The Lancet. Global health*, 7(6), e748–e760.
- Roser, M., Ortiz-Ospina, E., & Ritchie, H. (2013). Life Expectancy - Our World in Data. *Our World in Data*.
- Roser, M., & Ortiz-Ospina, E. (2016). Global Education - Our World in Data. *Our World in Data*.
- Rost, K., Smith, J. L., & Dickinson, M. (2004). The effect of improving primary care depression management on employee absenteeism and productivity. A randomized trial. *Medical Care*, 42(12), 1202–1210.
- Rowdhwal, S. S. S., & Chen, J. (2018). Toxic Effects of Di-2-ethylhexyl Phthalate: An Overview. *BioMed research international*, 2018, 1750368.
- Russell, A. E., Ford, T., Williams, R., & Russell, G. (2016). The association between socioeconomic disadvantage and attention deficit/hyperactivity disorder (ADHD): A systematic review. *Child psychiatry and human development*, 47(3), 440–458.
- Saffari, A., Silver, M. J., Zavattari, P., Moi, L., Columbano, A., Meaburn, E. L., & Dudbridge, F. (2018). Estimation of a significance threshold for epigenome-wide association studies. *Genetic Epidemiology*, 42(1), 20–33.

- Sahni, S., Talwar, A., Khanijo, S., & Talwar, A. (2017). Socioeconomic status and its relationship to chronic respiratory disease. *Advances in respiratory medicine*, 85(2), 97–108.
- Santor, D. A., Gregus, M., & Welch, A. (2006). FOCUS ARTICLE: eight decades of measurement in depression. *Measurement: Interdisciplinary Research & Perspective*, 4(3), 135–155.
- Sariaslan, A., Larsson, H., D’Onofrio, B., Långström, N., Fazel, S., & Lichtenstein, P. (2015). Does population density and neighborhood deprivation predict schizophrenia? A nationwide Swedish family-based study of 2.4 million individuals. *Schizophrenia Bulletin*, 41(2), 494–502.
- Sayer, N. A., Sackeim, H. A., Moeller, J. R., Prudic, J., Devanand, D. P., Coleman, E. A., & Kiersky, J. E. (1993). The relations between observer-rating and self-report of depressive symptomatology. *Psychological Assessment*, 5(3), 350–360.
- Schalkwyk, L. C., Pidsley, R., Wong, C. C. Y., & Touleimeimat, N. (2012). watermelon: Illumina 450 methylation array normalization and metrics.
- Seeman, T., Epel, E., Gruenewald, T., Karlamangla, A., & McEwen, B. S. (2010). Socio-economic differentials in peripheral biology: cumulative allostatic load. *Annals of the New York Academy of Sciences*, 1186, 223–239.
- Sethi, J. K., & Hotamisligil, G. S. (2021). Metabolic Messengers: tumour necrosis factor. *Nature Metabolism*, 3(10), 1302–1312.

- Shaw, M., Galobardes, B., Lawlor, D. A., Lynch, J., Wheeler, B., & Smith, G. D. (2007). Scottish index of multiple deprivation (SIMD) 2006. *The handbook of inequality and socioeconomic position: Concepts and measures* (pp. 137–139). Bristol University Press.
- Shen, X., Caramaschi, D., Adams, M. J., Walker, R. M., Min, J. L., Kwong, A., Hemani, G., et al. (2022). DNA methylome-wide association study of genetic risk for depression implicates antigen processing and immune responses. *Genome Medicine*, 14(1), 36.
- Shiina, T., Hosomichi, K., Inoko, H., & Kulski, J. K. (2009). The HLA genomic loci map: expression, interaction, diversity and disease. *Journal of Human Genetics*, 54(1), 15–39.
- Shiue, I. (2015). Urinary heavy metals, phthalates and polyaromatic hydrocarbons independent of health events are associated with adult depression: USA NHANES, 2011–2012. *Environmental Science and Pollution Research International*, 22(21), 17095–17103.
- Smith, B. H., Campbell, A., Linksted, P., Fitzpatrick, B., Jackson, C., Kerr, S. M., Deary, I. J., et al. (2013). Cohort Profile: Generation Scotland: Scottish Family Health Study (GS:SFHS). The study, its participants and their potential for genetic research on health and illness. *International Journal of Epidemiology*, 42(3), 689–700.
- Smith, C. A., & Holahan, M. R. (2014). Reduced hippocampal dendritic spine density and BDNF expression following acute postnatal exposure to di(2-ethylhexyl) phthalate in male Long Evans rats. *Plos One*, 9(10), e109522.
- Sobocki, P., Jönsson, B., Angst, J., & Rehnberg, C. (2006). Cost of depression in Europe. *The Journal of Mental Health Policy and Economics*, 9(2), 87–98.

- Spainhour, J. C., Lim, H. S., Yi, S. V., & Qiu, P. (2019). Correlation patterns between DNA methylation and gene expression in the cancer genome atlas. *Cancer informatics*, 18, 1176935119828776.
- Spitzer, R. L., Gibbon, M., Skodol, A. E., Williams, J. B. W., & First, M. B. (Eds.). (2002). *DSM-IV-TR Casebook: A Learning Companion to the Diagnostic and Statistical Manual of Mental Disorders*. Arlington, VA: American Psychiatric Publishing, Inc.
- Steffen, A., Nübel, J., Jacobi, F., Bätzing, J., & Holstiege, J. (2020). Mental and somatic comorbidity of depression: a comprehensive cross-sectional analysis of 202 diagnosis groups using German nationwide ambulatory claims data. *BMC Psychiatry*, 20(1), 142.
- Stetler, C., & Miller, G. E. (2011). Depression and hypothalamic-pituitary-adrenal activation: a quantitative summary of four decades of research. *Psychosomatic Medicine*, 73(2), 114–126.
- Stevenson, A. J., McCartney, D. L., Harris, S. E., Taylor, A. M., Redmond, P., Starr, J. M., Zhang, Q., et al. (2018). Trajectories of inflammatory biomarkers over the eighth decade and their associations with immune cell profiles and epigenetic ageing. *Clinical epigenetics*, 10(1), 159.
- Stevenson, A. J., McCartney, D. L., Hillary, R. F., Campbell, A., Morris, S. W., Bermingham, M. L., Walker, R. M., et al. (2020). Characterisation of an inflammation-related epigenetic score and its association with cognitive ability. *Clinical epigenetics*, 12(1), 113.
- Story Jovanova, O., Nedeljkovic, I., Spieler, D., Walker, R. M., Liu, C., Luciano, M., Bressler, J., et al. (2018). DNA Methylation Signatures of Depressive Symptoms in

- Middle-aged and Elderly Persons: Meta-analysis of Multiethnic Epigenome-wide Studies. *JAMA psychiatry*, 75(9), 949–959.
- Strawbridge, R., Young, A. H., & Cleare, A. J. (2017). Biomarkers for depression: recent insights, current challenges and future prospects. *Neuropsychiatric Disease and Treatment*, 13, 1245–1262.
- Stringhini, S., Carmeli, C., Jokela, M., Avendaño, M., Muennig, P., Guida, F., Ricceri, F., et al. (2017). Socioeconomic status and the 25 × 25 risk factors as determinants of premature mortality: a multicohort study and meta-analysis of 1·7 million men and women. *The Lancet*, 389(10075), 1229–1237.
- Sudak, D. M. (2012). Cognitive behavioral therapy for depression. *The Psychiatric Clinics of North America*, 35(1), 99–110.
- Suderman, M., Staley, J. R., French, R., Arathimos, R., Simpkin, A., & Tilling, K. (2018). dmrff: identifying differentially methylated regions efficiently with power and control. *BioRxiv*.
- Sullivan, P. F., de Geus, E. J. C., Willemsen, G., James, M. R., Smit, J. H., Zandbelt, T., Arolt, V., et al. (2009). Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. *Molecular Psychiatry*, 14(4), 359–375.
- Sullivan, P. F., Neale, M. C., & Kendler, K. S. (2000). Genetic epidemiology of major depression: review and meta-analysis. *The American Journal of Psychiatry*, 157(10), 1552–1562.
- Tabas, I., & Glass, C. K. (2013). Anti-inflammatory therapy in chronic disease: challenges and opportunities. *Science*, 339(6116), 166–172.

- Tallon, L. A., Manjourides, J., Pun, V. C., Salhi, C., & Suh, H. (2017). Cognitive impacts of ambient air pollution in the National Social Health and Aging Project (NSHAP) cohort. *Environment International*, 104, 102–109.
- Tang, D. D., & Gerlach, B. D. (2017). The roles and regulation of the actin cytoskeleton, intermediate filaments and microtubules in smooth muscle cell migration. *Respiratory Research*, 18(1), 54.
- Tost, J. (Ed.). (2009). *DNA Methylation*. Methods in Molecular Biology (Vol. 507). Totowa, NJ: Humana Press.
- Townsend, P. (1987). Deprivation. *Journal of social policy*, 16(02), 125–146.
- Turati, F., Galeone, C., Rota, M., Pelucchi, C., Negri, E., Bagnardi, V., Corrao, G., et al. (2014). Alcohol and liver cancer: a systematic review and meta-analysis of prospective studies. *Annals of Oncology*, 25(8), 1526–1535.
- Tyrrell, J., Mulugeta, A., Wood, A. R., Zhou, A., Beaumont, R. N., Tuke, M. A., Jones, S. E., et al. (2019). Using genetics to understand the causal influence of higher BMI on depression. *International Journal of Epidemiology*, 48(3), 834–848.
- Vazquez, C. E., & Cubbin, C. (2020). Socioeconomic Status and Childhood Obesity: a Review of Literature from the Past Decade to Inform Intervention Research. *Current obesity reports*, 9(4), 562–570.
- Vigo, D., Thornicroft, G., & Atun, R. (2016). Estimating the true global burden of mental illness. *The Lancet. Psychiatry*, 3(2), 171–178.

- Walker, J., Burke, K., Wanat, M., Fisher, R., Fielding, J., Mulick, A., Puntis, S., et al. (2018). The prevalence of depression in general hospital inpatients: a systematic review and meta-analysis of interview-based studies. *Psychological Medicine*, 48(14), 2285–2298.
- Wang, P. S., Simon, G. E., Avorn, J., Azocar, F., Ludman, E. J., McCulloch, J., Petukhova, M. Z., et al. (2007). Telephone screening, outreach, and care management for depressed workers and impact on clinical and work productivity outcomes: a randomized controlled trial. *The Journal of the American Medical Association*, 298(12), 1401–1411.
- Warren, J. R. (2009). Socioeconomic Status and Health across the Life Course: A Test of the Social Causation and Health Selection Hypotheses. *Social forces; a scientific medium of social study and interpretation*, 87(4), 2125–2153.
- Watanabe, K., Taskesen, E., van Bochoven, A., & Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. *Nature Communications*, 8(1), 1826.
- Watanabe, K., Umićević Mirkov, M., de Leeuw, C. A., van den Heuvel, M. P., & Posthuma, D. (2019). Genetic mapping of cell type specificity for complex traits. *Nature Communications*, 10(1), 3222.
- Werner, S., Malaspina, D., & Rabinowitz, J. (2007). Socioeconomic status at birth is associated with risk of schizophrenia: population-based multilevel study. *Schizophrenia Bulletin*, 33(6), 1373–1378.
- Wickramarachchi, D. C., Theofilopoulos, A. N., & Kono, D. H. (2010). Immune pathology associated with altered actin cytoskeleton regulation. *Autoimmunity*, 43(1), 64–75.

- Wilson, S., Hicks, B. M., Foster, K. T., McGue, M., & Iacono, W. G. (2015). Age of onset and course of major depressive disorder: associations with psychosocial functioning outcomes in adulthood. *Psychological Medicine*, 45(3), 505–514.
- Wimalasena, N. N., Chang-Richards, A., Wang, K. I.-K., & Dirks, K. N. (2021). Housing Risk Factors Associated with Respiratory Disease: A Systematic Review. *International Journal of Environmental Research and Public Health*, 18(6).
- Wong, M. C. S., Lao, X. Q., Ho, K.-F., Goggins, W. B., & Tse, S. L. A. (2017). Incidence and mortality of lung cancer: global trends and association with socioeconomic status. *Scientific Reports*, 7(1), 14300.
- World Health Organization. (2017). *Depression and Other Common Mental Disorders. Global Health Estimates*. Geneva: World Health Organization; 2017. Licence: CC BY-NC-SA 3.0 IGO.
- Wray, Naomi R, Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., Adams, M. J., et al. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*, 50(5), 668–681.
- Wray, N R, Pergadia, M. L., Blackwood, D. H. R., Penninx, B. W. J. H., Gordon, S. D., Nyholt, D. R., Ripke, S., et al. (2012). Genome-wide association study of major depressive disorder: new results, meta-analysis, and lessons learned. *Molecular Psychiatry*, 17(1), 36–48.
- W H O. (2021). WHO. Retrieved May 12, 2022, from <https://www.who.int/news-room/fact-sheets/detail/depression>

- Xue, A., Jiang, L., Zhu, Z., Wray, N. R., Visscher, P. M., Zeng, J., & Yang, J. (2020). Genome-wide analyses of behavioural traits biased by misreports and longitudinal changes. *medRxiv*.
- Yang, T., Nie, Z., Shu, H., Kuang, Y., Chen, X., Cheng, J., Yu, S., et al. (2020). The role of BDNF on neural plasticity in depression. *Frontiers in Cellular Neuroscience*, 14, 82.
- Zhou, W., Laird, P. W., & Shen, H. (2017). Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucleic Acids Research*, 45(4), e22.
- Zimmerman, M., Ellison, W., Young, D., Chelminski, I., & Dalrymple, K. (2015). How many different ways do patients meet the diagnostic criteria for major depressive disorder? *Comprehensive Psychiatry*, 56, 29–34.
- Zisook, S., Lesser, I., Stewart, J. W., Wisniewski, S. R., Balasubramani, G. K., Fava, M., Gilmer, W. S., et al. (2007). Effect of age at onset on the course of major depressive disorder. *The American Journal of Psychiatry*, 164(10), 1539–1546.

Systematic Review References

- Adewuya, A. O., Ola, B. A., Ajayi, O. E., Oyedeji, A. O., Balogun, M. O., & Mosaku, S. K. (2006). Prevalence and correlates of major depressive disorder in Nigerian outpatients with heart failure. *Psychosomatics*, 47(6), 479–485.
- Ahmadi, H., Montaser-Kouhsari, L., Nowroozi, M. R., & Bazargan-Hejazi, S. (2011). Male infertility and depression: a neglected problem in the Middle East. *The Journal of Sexual Medicine*, 8(3), 824–830.

Alamri, S. H., Bari, A. I., & Ali, A. T. (2017). Depression and associated factors in hospitalized elderly: a cross-sectional study in a Saudi teaching hospital. *Annals of Saudi Medicine*, 37(2), 122–129.

AlBekairy, A., AbuRuz, S., Alsabani, B., Alshehri, A., Aldebasi, T., Alkatheri, A., & Almodaimegh, H. (2017). Exploring Factors Associated with Depression and Anxiety among Hospitalized Patients with Type 2 Diabetes Mellitus. *Medical principles and practice : international journal of the Kuwait University, Health Science Centre*, 26(6), 547–553.

Almeida, O. P., Draper, B., Pirkis, J., Snowdon, J., Lautenschlager, N. T., Byrne, G., Sim, M., et al. (2012). Anxiety, depression, and comorbid anxiety and depression: risk factors and outcome over two years. *International Psychogeriatrics*, 24(10), 1622–1632.

Almeida-Filho, N., Lessa, I., Magalhães, L., Araújo, M. J., Aquino, E., James, S. A., & Kawachi, I. (2004). Social inequality and depressive disorders in Bahia, Brazil: interactions of gender, ethnicity, and social class. *Social Science & Medicine*, 59(7), 1339–1353.

Alonso-Morán, E., Satylganova, A., Orueta, J. F., & Nuño-Solinis, R. (2014). Prevalence of depression in adults with type 2 diabetes in the Basque Country: relationship with glycaemic control and health care costs. *BMC Public Health*, 14, 769.

Alqahtani, M. S., Alshbriqe, A. A., Awwadh, A. A., Alyami, T. A., Alshomrani, M. S., & Alhazzani, A. (2018). Prevalence and Risk Factors for Depression among Caregivers of Alzheimer's Disease Patients in Saudi Arabia. *Neurology research international*, 2018, 2501835.

Altino, D. M., Nogueira-Martins, L. A., de Barros, A. L. B. L., & Lopes, J. de L. (2017). Predictive Factors of Anxiety and Depression in Patients with Acute Coronary Syndrome. *Archives of psychiatric nursing*, 31(6), 549–552.

Ambugo, E. A. (2014). Cross-country variation in the sociodemographic factors associated with major depressive episode in Norway, the United Kingdom, Ghana, and Kenya. *Social Science & Medicine*, 113, 154–160.

Andersen, I., Thielen, K., Bech, P., Nygaard, E., & Diderichsen, F. (2011). Increasing prevalence of depression from 2000 to 2006. *Scandinavian Journal of Public Health*, 39(8), 857–863.

Andersson, D., Magnusson, H., Carstensen, J., & Borgquist, L. (2011). Co-morbidity and health care utilisation five years prior to diagnosis for depression. A register-based study in a Swedish population. *BMC Public Health*, 11, 552.

Asangbeh, S. L., Sobngwi, J. L., Ekali, G. L., Eyoum, C., & Msellati, P. (2016). Predictors of depression among patients on art in a rural health district in North West Cameroon. *AIDS Care*, 28(2), 205–208.

Aslund, C., Nilsson, K. W., Starrin, B., & Sjöberg, R. L. (2007). Shaming experiences and the association between adolescent depression and psychosocial risk factors. *European Child & Adolescent Psychiatry*, 16(5), 298–304.

Assari, S., Lankarani, M. M., & Caldwell, C. H. (2018). Does Discrimination Explain High Risk of Depression among High-Income African American Men? *Behavioral sciences (Basel, Switzerland)*, 8(4).

Avanci, J., Assis, S., Oliveira, R., & Pires, T. (2012). Childhood depression. Exploring the association between family violence and other psychosocial factors in low-income Brazilian schoolchildren. *Child and Adolescent Psychiatry and Mental Health Vol 6 2012*, ArtID 26, 6.

Ayazi, T., Lien, L., Eide, A. H., Ruom, M. M., & Hauff, E. (2012). What are the risk factors for the comorbidity of posttraumatic stress disorder and depression in a war-affected population? A cross-sectional community study in South Sudan. *BMC Psychiatry Vol 12 2012*, ArtID 175, 12.

Balci, G., Oter, G., Akdag, H., Bekki, A., Kisa, C., & Goka, E. (2016). Factors associated with depression in patients with schizophrenia. *Journal of Mood Disorders*, 6(2), 54.

Batterham, P. J., Christensen, H., & Mackinnon, A. J. (2009). Modifiable risk factors predicting major depressive disorder at four year follow-up: a decision tree approach. *BMC Psychiatry*, 9, 75.

Bautista, R. E. D., & Erwin, P. A. (2013). Analyzing depression coping strategies of patients with epilepsy: a preliminary study. *Seizure*, 22(9), 686–691.

Beard, J. R., Tracy, M., Vlahov, D., & Galea, S. (2008). Trajectory and socioeconomic predictors of depression in a prospective study of residents of New York City. *Annals of Epidemiology*, 18(3), 235–243.

Bekaroğlu, M., Uluutku, N., Tanriöver, S., & Kirpinar, I. (1991). Depression in an elderly population in Turkey. *Acta Psychiatrica Scandinavica*, 84(2), 174–178.

Berganza, C. E., & Aguilar, G. (1992). Depression in Guatemalan adolescents. *Adolescence*, 27(108), 771–82.

Bhatia, R., Hartman, C., Kallen, M. A., Graham, J., & Giordano, T. P. (2011). Persons newly diagnosed with HIV infection are at high risk for depression and poor linkage to care: results from the Steps Study. *AIDS and Behavior*, 15(6), 1161–1170.

Blanco, C., Okuda, M., Markowitz, J. C., Liu, S.-M., Grant, B. F., & Hasin, D. S. (2010). The epidemiology of chronic major depressive disorder and dysthymic disorder: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *The Journal of Clinical Psychiatry*, 71(12), 1645–1656.

Blay, S. L., Fillenbaum, G. G., Marinho, V., Andreoli, S. B., & Gastal, F. L. (2011). Increased health burden associated with comorbid depression in older Brazilians with diabetes. *Journal of Affective Disorders*, 134(1–3), 77–84.

Braich, P. S., Lal, V., Hollands, S., & Almeida, D. R. (2012). Burden and depression in the caregivers of blind patients in India. *Ophthalmology*, 119(2), 221–226.

Butterworth, P., Olesen, S. C., & Leach, L. S. (2012). The role of hardship in the association between socio-economic position and depression. *The Australian and New Zealand Journal of Psychiatry*, 46(4), 364–373.

Buvneshkumar, M., John, K. R., & Logaraj, M. (2018). A study on prevalence of depression and associated risk factors among elderly in a rural block of Tamil Nadu. *Indian journal of public health*, 62(2), 89–94.

Çakıcı, M., Gökçe, Ö., Babayiğit, A., Çakıcı, E., & Eş, A. (2017). Depression: point-prevalence and risk factors in a North Cyprus household adult cross-sectional study. *BMC Psychiatry*, 17(1), 387.

Castilla-Puentes, R. C., Secin, R., Grau, A., Galeno, R., Feijo de Mello, M., Pena, N., & Sanchez-Russi, C. A. (2008). A multicenter study of major depressive disorder among emergency department patients in Latin-American countries. *Depression and Anxiety*, 25(12), E199-204.

Castro, L. S., Castro, J., Hoexter, M. Q., Quarantini, L. C., Kauati, A., Mello, L. E., Santos-Silva, R., et al. (2013). Depressive symptoms and sleep: a population-based polysomnographic study. *Psychiatry Research*, 210(3), 906–912.

Chang, B., Steimel, J., Moller, D. R., Baughman, R. P., Judson, M. A., Yeager, H., Teirstein, A. S., et al. (2001). Depression in sarcoidosis. *American Journal of Respiratory and Critical Care Medicine*, 163(2), 329–334.

Chiavegatto Filho, A. D. P., Kawachi, I., Wang, Y. P., Viana, M. C., & Andrade, L. H. S. G. (2013). Does income inequality get under the skin? A multilevel analysis of depression, anxiety and mental disorders in Sao Paulo, Brazil. *Journal of Epidemiology and Community Health*, 67(11), 966–972.

Clayton, P. J., Halikas, J. A., & Maurice, W. L. (1972). The depression of widowhood. *The British Journal of Psychiatry*, 120(554), 71–77.

Cole, A. L., Lee, P. J., Hughes, D. A., Deegan, P. B., Waldek, S., & Lachmann, R. H. (2007). Depression in adults with Fabry disease: a common and under-diagnosed problem. *Journal of Inherited Metabolic Disease*, 30(6), 943–951.

Comstock, G. W., & Helsing, K. J. (1976). Symptoms of depression in two communities. *Psychological Medicine*, 6(4), 551–563.

Coombe, C. M. (2008). The effects of urban residential environments on mental well being: A multilevel analysis of neighborhood stability, middle income composition and depression in detroit. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 68(10-B), 6613.

Deveci, S. E., Açık, Y., Dag, D. G., Tokdemir, M., & Gündoğdu, C. (2010). The frequency of depression and menopause-related symptoms in postmenopausal women living in a province in Eastern Turkey, and the factors that affect depressive status. *Medical Science Monitor*, 16(4), PH40-47.

Dillard, D. A., Smith, J. J., Ferucci, E. D., & Lanier, A. P. (2012). Depression prevalence and associated factors among Alaska Native people: the Alaska education and research toward health (EARTH) study. *Journal of Affective Disorders*, 136(3), 1088–1097.

Dismuke, C. E., & Egede, L. E. (2010). Association between major depression, depressive symptoms and personal income in US adults with diabetes. *General Hospital Psychiatry*, 32(5), 484–491.

Donoho, C. J., LeardMann, C., O'Malley, C. A., Walter, K. H., Riviere, L. A., Curry, J. F., & Adler, A. B. (2018). Depression among military spouses: Demographic, military, and service member psychological health risk factors. *Depression and Anxiety*, 35(12), 1137–1144.

Duthé, G., Rossier, C., Bonnet, D., Soura, A. B., & Corker, J. (2016). Mental health and urban living in sub-Saharan Africa: major depressive episodes among the urban poor in Ouagadougou, Burkina Faso. *Population health metrics*, 14(1), 18.

Feuillet, P., Lert, F., Tron, L., Aubriere, C., Spire, B., Dray-Spira, R., & Agence Nationale de Recherche sur le Sida et les Hépatites Virales (ANRS)-VIH: Enquête sur les personnes atteintes (Vespa2) Study Group. (2017). Prevalence of and factors associated with depression among people living with HIV in France. *HIV Medicine*, 18(6), 383–394.

Figueiredo, J. H. C., Silva, N. A. de S. E., Pereira, B. de B., & Oliveira, G. M. M. de. (2017). Major Depression and Acute Coronary Syndrome-Related Factors. *Arquivos brasileiros de cardiologia*, 108(3), 217–227.

Freeman, A., Tyrovolas, S., Koyanagi, A., Chatterji, S., Leonardi, M., Ayuso-Mateos, J. L., Tobiasz-Adamczyk, B., et al. (2016). The role of socio-economic status in depression: results from the COURAGE (aging survey in Europe). *BMC Public Health*, 16(1), 1098.

Goodman, E., Slap, G. B., & Huang, B. (2003). The public health impact of socioeconomic status on adolescent depression and obesity. *American Journal of Public Health*, 93(11), 1844–1850.

Gupta, R., Dandu, M., Packel, L., Rutherford, G., Leiter, K., Phaladze, N., Korte, F. P., et al. (2010). Depression and HIV in Botswana: a population-based study on gender-specific socioeconomic and behavioral correlates. *Plos One*, 5(12), e14252.

Habtewold, T. D., Alemu, S. M., & Haile, Y. G. (2016). Sociodemographic, clinical, and psychosocial factors associated with depression among type 2 diabetic outpatients in Black Lion General Specialized Hospital, Addis Ababa, Ethiopia: a cross-sectional study. *BMC Psychiatry*, 16, 103.

Hamano, T., Li, X., Lönn, S. L., Nabika, T., Sundquist, J., & Sundquist, K. (2018). Is familial risk for depression confounded by individual and familial socioeconomic factors and neighborhood environmental factors? A 7-year follow-up study in Sweden. *Psychiatry Research*, 266, 30–35.

Hankerson, S. H., Lee, Y. A., Brawley, D. K., Braswell, K., Wickramaratne, P. J., & Weissman, M. M. (2015). Screening for Depression in African-American Churches. *American Journal of Preventive Medicine*, 49(4), 526–533.

Hansen, T., Slagsvold, B., & Veenstra, M. (2017). Educational inequalities in late-life depression across Europe: results from the generations and gender survey. *European journal of ageing*, 14(4), 407–418.

Ho, H. C., Lau, K. K.-L., Yu, R., Wang, D., Woo, J., Kwok, T. C. Y., & Ng, E. (2017). Spatial Variability of Geriatric Depression Risk in a High-Density City: A Data-Driven Socio-Environmental Vulnerability Mapping Approach. *International Journal of Environmental Research and Public Health*, 14(9).

Hu, A., Xue, Z., Mwansisya, T. E., Zhou, A., Pu, W., Chen, X., Sun, M., et al. (2015). Major depressive disorder in hemodialysis patients in China. *Asia-Pacific psychiatry : official journal of the Pacific Rim College of Psychiatrists*, 7(1), 78–84.

Jaffary, M., Samo, J. A., & Mujtaba, S. W. A. (2016). Rate of Anxiety and Depression in Cases of Bronchial Asthma. *Pakistan Journal of Medical & Health Sciences*, 10(4), 1353–1355.

Jayashree, K., Mithra, P. P., Nair, M. K. C., Unnikrishnan, B., & Pai, K. (2018). Depression and Anxiety Disorders among Schoolgoing Adolescents in an Urban Area of South India. *Indian journal of community medicine : official publication of Indian Association of Preventive & Social Medicine*, 43(Suppl 1), S28–S32.

Joergensen, T. S. H., Maartensson, S., Ibfelt, E. H., Joergensen, M. B., Wium-Andersen, I. K., Wium-Andersen, M. K., Prescott, E., et al. (2016). Depression following acute coronary syndrome: a Danish nationwide study of potential risk factors. *Social Psychiatry and Psychiatric Epidemiology*, 51(11), 1509–1523.

Johannsen, A., Rydmark, I., Söder, B., & Asberg, M. (2007). Gingival inflammation, increased periodontal pocket depth and elevated interleukin-6 in gingival crevicular fluid of depressed women on long-term sick leave. *Journal of periodontal research*, 42(6), 546–552.

Johnston, E., Johnson, S., McLeod, P., & Johnston, M. (2004). The relation of body mass index to depressive symptoms. *Canadian Journal of Public Health*, 95(3), 179–183.

Khalil, M. I., Rahman, M. R., Munira, S., & Jahan, M. U. (2018). Risk factors of major depressive disorder in parkinson's disease. *Bangladesh Medical Research Council bulletin*, 44(1), 9.

Khazaeipour, Z., Taheri-Otaghsara, S.-M., & Naghdi, M. (2015). Depression following spinal cord injury: its relationship to demographic and socioeconomic indicators. *Topics in spinal cord injury rehabilitation*, 21(2), 149–155.

Kim, J.-M., Shin, I.-S., Yoon, J.-S., & Stewart, R. (2002). Prevalence and correlates of late-life depression compared between urban and rural populations in Korea. *International Journal of Geriatric Psychiatry*, 17(5), 409–415.

Kim, K. W., Kim, S. H., Shin, J. H., Choi, B. Y., Nam, J. H., & Park, S.-C. (2014). Psychosocial, physical, and autonomic correlates of depression in korean adults: results from a county-based depression screening study. *Psychiatry investigation*, 11(4), 402–411.

Kim, M. C., Jung, Y. S., Song, Y. S., Lee, J. I., Park, J. H., Sohn, C. I., Choi, K. Y., et al. (2016). Factors Associated with Anxiety and Depression in Korean Patients with Inactive Inflammatory Bowel Disease. *Gut and Liver*, 10(3), 399–405.

Kinyanda, E., Kuteesa, M., Scholten, F., Mugisha, J., Baisley, K., & Seeley, J. (2016). Risk of major depressive disorder among older persons living in HIV-endemic central and southwestern Uganda. *AIDS Care*, 28(12), 1516–1521.

Kinyanda, E., Woodburn, P., Tugumisirize, J., Kagugube, J., Ndyabangi, S., & Patel, V. (2011). Poverty, life events and the risk for depression in Uganda. *Social Psychiatry and Psychiatric Epidemiology*, 46(1), 35–44.

Kollannoor-Samuel, G., Wagner, J., Damio, G., Segura-Pérez, S., Chhabra, J., Vega-López, S., & Pérez-Escamilla, R. (2011). Social support modifies the association between household food insecurity and depression among Latinos with uncontrolled type 2 diabetes. *Journal of Immigrant and Minority Health*, 13(6), 982–989.

Kosidou, K., Dalman, C., Lundberg, M., Hallqvist, J., Isacson, G., & Magnusson, C. (2011). Socioeconomic status and risk of psychological distress and depression in the Stockholm Public Health Cohort: a population-based study. *Journal of Affective Disorders*, 134(1–3), 160–167.

Kui, C., Yingfu, P., Chenling, X., Wenqing, W., Xiuhua, L., & Di, S. (2014). What are the predictors of major depression in adult patients with epilepsy? *Epileptic disorders : international epilepsy journal with videotape*, 16(1), 74–79.

Kumar, A., Clark, S., Boudreaux, E. D., & Camargo, C. A. (2004). A multicenter study of depression among emergency department patients. *Academic Emergency Medicine*, 11(12), 1284–1289.

Lee, C.-T., Chiang, Y.-C., Huang, J.-Y., Tantoh, D. M., Nfor, O. N., Lee, J.-F., Chang, C.-C., et al. (2016). Incidence of Major Depressive Disorder: Variation by Age and Sex in Low-Income Individuals: A Population-Based 10-Year Follow-Up Study. *Medicine*, 95(15), e3110.

Lemogne, C., Meneton, P., Wiernik, E., Quesnot, A., Consoli, S. M., Ducimetière, P., Nabi, H., et al. (2017). When Blue-Collars Feel Blue: Depression and Low Occupational Grade as Synergistic Predictors of Incident Cardiac Events in Middle-Aged Working Individuals. *Circulation. Cardiovascular Quality and Outcomes*, 10(2).

Lin, F. G., Hsieh, Y. H., & Tung, H. J. (2012). Interactive effects of family socioeconomic status and body mass index on depression in school-aged children, 21(1), 64–72.

Long, M. D., Kappelman, M. D., Martin, C. F., Chen, W., Anton, K., & Sandler, R. S. (2014). Risk factors for depression in the elderly inflammatory bowel disease population. *Journal of Crohn's & colitis*, 8(2), 113–119.

Lotfaliany, M., Bowe, S. J., Kowal, P., Orellana, L., Berk, M., & Mohebbi, M. (2018). Depression and chronic diseases: Co-occurrence and communality of risk factors. *Journal of Affective Disorders*, 241, 461–468.

- Lu, J., Ruan, Y., Huang, Y., Yao, J., Dang, W., & Gao, C. (2008). Major depression in Kunming: prevalence, correlates and co-morbidity in a south-western city of China. *Journal of Affective Disorders*, 111(2–3), 221–226.
- Madianos, M., Economou, M., Alexiou, T., & Stefanis, C. (2011). Depression and economic hardship across Greece in 2008 and 2009: two cross-sectional surveys nationwide. *Social Psychiatry and Psychiatric Epidemiology*, 46(10), 943–952.
- Magklara, K., Bellos, S., Niakas, D., Stylianidis, S., Kolaitis, G., Mavreas, V., & Skapinakis, P. (2015). Depression in late adolescence: a cross-sectional study in senior high schools in Greece. *BMC Psychiatry*, 15, 199.
- Maharaj, R G, Reid, S. D., Misir, A., & Simeon, D. T. (2005). Depression and its associated factors among patients attending chronic disease clinics in Southwest Trinidad. *West Indian Medical Journal*, 54(6), 369–374.
- Maharaj, Rohan G. (2007). Depression and the nature of Trinidadian family practice: a cross-sectional study. *BMC family practice*, 8, 25.
- Mahmood, S., Hassan, S. Z., Tabraze, M., Khan, M. O., Javed, I., Ahmed, A., Siddiqui, O. M., et al. (2017). Prevalence and predictors of depression amongst hypertensive individuals in karachi, pakistan. *Cureus*, 9(6), e1397.
- Mccormick, N., Trupin, L., Yelin, E. H., & Katz, P. P. (2018). Socioeconomic predictors of incident depression in systemic lupus erythematosus. *Arthritis Care & Research*, 70(1), 104–113.
- McDonald, L. S. (2011). The relationship between joblessness and depression and depressive symptomatology among a working age adult cohort in post-conflict Bosnia-Herzegovina (2001-2004). *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 72(3-B), 1445.

Modabber-Nia, M.-J., Shodjai-Tehrani, H., Moosavi, S.-R., Jahanbakhsh-Asli, N., & Fallahi, M. (2007). The prevalence of depression among high school and preuniversity adolescents: Rasht, northern Iran. *Archives of Iranian medicine*, 10(2), 141–146.

Mojtabai, R., & Olfson, M. (2004). Major depression in community-dwelling middle-aged and older adults: prevalence and 2- and 4-year follow-up symptoms. *Psychological Medicine*, 34(4), 623–634.

Muhwezi, W. W., Agren, H., Neema, S., Maganda, A. K., & Musisi, S. (2008). Life events associated with major depression in Ugandan primary healthcare (PHC) patients: issues of cultural specificity. *The International Journal of Social Psychiatry*, 54(2), 144–163.

Murata, C., Kondo, K., Hirai, H., Ichida, Y., & Ojima, T. (2008). Association between depression and socio-economic status among community-dwelling elderly in Japan: the Aichi Gerontological Evaluation Study (AGES). *Health & Place*, 14(3), 406–414.

Murcia, M., Chastang, J.-F., & Niedhammer, I. (2015). Educational inequalities in major depressive and generalized anxiety disorders: results from the French national SIP study. *Social Psychiatry and Psychiatric Epidemiology*, 50(6), 919–928.

Mushtaque, A., Gulati, R., Hossain, M. M., & Azmi, S. A. (2016). Prevalence of depression in patients of type 2 diabetes mellitus: A cross sectional study in a tertiary care centre. *Diabetes & metabolic syndrome*, 10(4), 238–241.

Nahon, S., Lahmek, P., Durance, C., Olympie, A., Lesgourgues, B., Colombel, J.-F., & Gendre, J.-P. (2012). Risk factors of anxiety and depression in inflammatory bowel disease. *Inflammatory Bowel Diseases*, 18(11), 2086–2091.

Nicholson, A., Pikhart, H., Pajak, A., Malyutina, S., Kubinova, R., Peasey, A., Topor-Madry, R., et al. (2008). Socio-economic status over the life-course and depressive symptoms in men and women in Eastern Europe. *Journal of Affective Disorders*, 105(1–3), 125–136.

Nik Jaafar, N. R., Selamat Din, S. H., Mohamed Saini, S., Ahmad, S. N. A., Midin, M., Sidi, H., Silim, U. A., et al. (2014). Clinical depression while caring for loved ones with breast cancer. *Comprehensive Psychiatry*, 55 Suppl 1, S52-9.

Nikulina, V., Widom, C. S., & Czaja, S. (2011). The role of childhood neglect and childhood poverty in predicting mental health, academic achievement and crime in adulthood. *American Journal of Community Psychology*, 48(3–4), 309–321.

Nomoto, S. H., Longhi, R. M. P., de Barros, B. P., Croda, J., Ziff, E. B., & Castelon Konkiewitz, E. (2015). Socioeconomic disadvantage increasing risk for depression among recently diagnosed HIV patients in an urban area in Brazil: cross-sectional study. *AIDS Care*, 27(8), 979–985.

Ompad, D. C., Palamar, J. J., Krause, K. D., Kapadia, F., & Halkitis, P. N. (2018). Reliability and validity of a material resources scale and its association with depression among young men who have sex with men: the P18 cohort study. *American journal of men's health*, 12(5), 1384–1397.

Pahkala, K., Kesti, E., Köngäs-Saviaro, P., Laippala, P., & Kivelä, S. L. (1995). Prevalence of depression in an aged population in Finland. *Social Psychiatry and Psychiatric Epidemiology*, 30(3), 99–106.

Park, J. H., Kim, K. W., Kim, M.-H., Kim, M. D., Kim, B.-J., Kim, S.-K., Kim, J. L., et al. (2012). A nationwide survey on the prevalence and risk factors of late life depression in South Korea. *Journal of Affective Disorders*, 138(1–2), 34–40.

Pikhartova, J., Chandola, T., Kubinova, R., Bobak, M., Nicholson, A., & Pikhart, H. (2009). Neighbourhood socioeconomic indicators and depressive symptoms in the Czech Republic: a population based study. *International journal of public health*, 54(4), 283–293.

Poongothai, S., Pradeepa, R., Ganesan, A., & Mohan, V. (2009). Prevalence of depression in a large urban South Indian population--the Chennai Urban Rural Epidemiology Study (CURES-70). *Plos One*, 4(9), e7185.

Post, D. M., Gehlert, S., Hade, E. M., Reiter, P. L., Ruffin, M., & Paskett, E. D. (2013). Depression and SES in women from Appalachia. *Journal of Rural Mental Health*, 37(1), 2–15.

Qiu, J., Yang, M., Chen, W., Gao, X., Liu, S., Shi, S., & Xie, B. (2012). Prevalence and correlates of major depressive disorder in breast cancer survivors in Shanghai, China. *Psycho-Oncology*, 21(12), 1331–1337.

Raji, S. O., Lawani, A. O., & James, B. O. (2016). Prevalence and correlates of major depression among Nigerian adults with sickle cell disease. *International Journal of Psychiatry in Medicine*, 51(5), 456–466.

Reinherz, H. Z., Giaconia, R. M., Lefkowitz, E. S., Pakiz, B., & Frost, A. K. (1993). Prevalence of psychiatric disorders in a community population of older adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, 32(2), 369–377.

Roze, M., Vandentorren, S., van der Waerden, J., & Melchior, M. (2018). Factors associated with depression among homeless mothers. Results of the ENFAMS survey. *Journal of Affective Disorders*, 229, 314–321.

Saxena, A., & Suman, A. (2015). Magnitude and determinants of depression in acute stroke patients admitted in a rural tertiary care hospital. *Journal of neurosciences in rural practice*, 6(2), 202–207.

Shi, J., Zhang, Y., Liu, F., Li, Y., Wang, J., Flint, J., Gao, J., et al. (2014). Associations of educational attainment, occupation, social class and major depressive disorder among Han Chinese women. *Plos One*, 9(1), e86674.

Sidik, S. M., Arroll, B., Goodyear-Smith, F., & Ahmad, R. (2012). Prevalence of depression among women attending a primary urban care clinic in Malaysia. *Singapore Medical Journal*, 53(7), 468–473.

Smith, K. L. W., Matheson, F. I., Moineddin, R., & Glazier, R. H. (2007). Gender, income and immigration differences in depression in canadian urban centres. *Canadian Journal of Public Health*, 98(2), 149–153.

Spendlove, D. C., West, D. W., & Stanish, W. M. (1984). Risk factors and the prevalence of depression in Mormon women. *Social Science & Medicine*, 18(6), 491–495.

Stansfeld, S. A., Clark, C., Rodgers, B., Caldwell, T., & Power, C. (2008). Childhood and adulthood socio-economic position and midlife depressive and anxiety disorders. *The British Journal of Psychiatry*, 192(2), 152–153.

Stolic, R. V., Trajkovic, G. Z., Mihailovic, B., Sipic, M. V., Celic, D. B., Lazic, S. F., Nikolic, G. R., et al. (2010). Characteristics of depression in obese people living in an insecure environment. *Indian Journal of Medical Sciences*, 64(7), 307–314.

Tani, Y., Fujiwara, T., Kondo, N., Noma, H., Sasaki, Y., & Kondo, K. (2016). Childhood Socioeconomic Status and Onset of Depression among Japanese Older Adults: The JAGES Prospective Cohort Study. *The American Journal of Geriatric Psychiatry*, 24(9), 717–726.

Tilahune, A. B., Bekele, G., Mekonnen, N., & Tamiru, E. (2016). Prevalence of unrecognized depression and associated factors among patients attending medical outpatient department in Adare Hospital, Hawassa, Ethiopia. *Neuropsychiatric Disease and Treatment*, 12, 2723–2729.

Tintle, N., Bacon, B., Kostyuchenko, S., Gutkovich, Z., & Bromet, E. J. (2011). Depression and its correlates in older adults in Ukraine. *International Journal of Geriatric Psychiatry*, 26(12), 1292–1299.

Turner, R. J., Lloyd, D. A., & Roszell, P. (1999). Personal resources and the social distribution of depression. *American Journal of Community Psychology*, 27(5), 643–672.

Vallée, J., Cadot, E., Roustit, C., Parizot, I., & Chauvin, P. (2011). The role of daily mobility in mental health inequalities: the interactive influence of activity space and neighbourhood of residence on depression. *Social Science & Medicine*, 73(8), 1133–1144.

- Verger, P., Guagliardo, V., Gilbert, F., Rouillon, F., & Kovess-Masfety, V. (2010). Psychiatric disorders in students in six French universities: 12-month prevalence, comorbidity, impairment and help-seeking. *Social Psychiatry and Psychiatric Epidemiology*, 45(2), 189–199.
- Wainwright, N. W. J., & Surtees, P. G. (2004). Area and individual circumstances and mood disorder prevalence. *The British Journal of Psychiatry*, 185, 227–232.
- Walters, K., Breeze, E., Wilkinson, P., Price, G. M., Bulpitt, C. J., & Fletcher, A. (2004). Local area deprivation and urban-rural differences in anxiety and depression among people older than 75 years in Britain. *American Journal of Public Health*, 94(10), 1768–1774.
- Wang, J. J. (2001). Prevalence and correlates of depressive symptoms in the elderly of rural communities in southern Taiwan. *The journal of nursing research : JNR*, 9(3), 1–12.
- Wang, J., Patten, S. B., Currie, S., Sareen, J., & Schmitz, N. (2012). A population-based longitudinal study on work environmental factors and the risk of major depressive disorder. *American Journal of Epidemiology*, 176(1), 52–59.
- Whelan-Goodinson, R., Ponsford, J. L., Schönberger, M., & Johnston, L. (2010). Predictors of psychiatric disorders following traumatic brain injury. *The Journal of head trauma rehabilitation*, 25(5), 320–329.
- Whooley, M. A., Kiefe, C. I., Chesney, M. A., Markovitz, J. H., Matthews, K., Hulley, S. B., & CARDIA Study. (2002). Depressive symptoms, unemployment, and loss of income: The CARDIA Study. *Archives of Internal Medicine*, 162(22), 2614–2620.
- Wilson, K. C., Chen, R., Taylor, S., McCracken, C. F., & Copeland, J. R. (1999). Socio-economic deprivation and the prevalence and prediction of depression in older community residents. The MRC-ALPHA Study. *The British Journal of Psychiatry*, 175, 549–553.
- Wirback, T., Möller, J., Larsson, J.-O., & Engström, K. (2018). Social differences in diagnosed depression among adolescents in a Swedish population based cohort. *BMC Psychiatry*, 18(1), 216.

Xu, X., Zhao, X., Qian, D., Dong, Q., & Gu, Z. (2015). Investigating Factors Associated with Depression of Type 2 Diabetic Retinopathy Patients in China. *Plos One*, 10(7), e0132616.

Yaka, E., Keskinoglu, P., Ucku, R., Yener, G. G., & Tunca, Z. (2014). Prevalence and risk factors of depression among community dwelling elderly. *Archives of Gerontology and Geriatrics*, 59(1), 150–154.

Yen, P. N., Lin, I.-F., Chang, W. P., Wang, J.-D., Chang, T.-C., Kuo, K.-L., Hwang, J.-S., et al. (2014). Risk factors of depression after prolonged low-dose rate environmental radiation exposure. *International journal of radiation biology*, 90(10), 859–866.

Yu, S., Yang, H., Guo, X., Zheng, L., & Sun, Y. (2016). Prevalence of Depression among Rural Residents with Diabetes Mellitus: A Cross-Sectional Study from Northeast China. *International Journal of Environmental Research and Public Health*, 13(6).

Yusim, A., Anbarasan, D., Hall, B., Goetz, R., Neugebauer, R., Stewart, T., Abou, J., et al. (2010). Sociocultural domains of depression among indigenous populations in Latin America. *International Review of Psychiatry*, 22(4), 370–377.

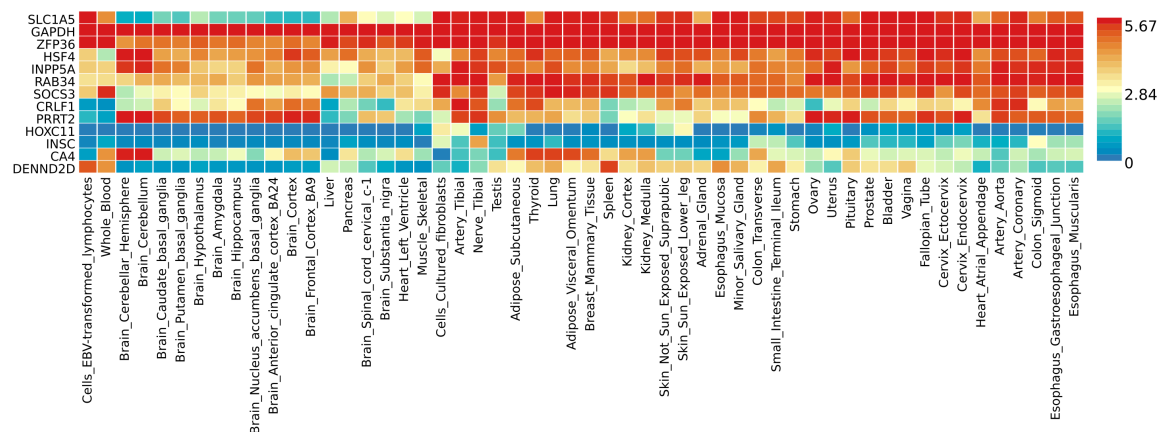
Zhao, D., Hu, C., Chen, J., Dong, B., Ren, Q., Yu, D., Zhao, Y., et al. (2018). Risk factors of geriatric depression in rural China based on a generalized estimating equation. *International Psychogeriatrics*, 30(10), 1489–1497.

Zhong, B., Chen, H., Zhang, J., Xu, H., Zhou, C., Yang, F., Song, J., et al. (2010). Prevalence, correlates and recognition of depression among inpatients of general hospitals in Wuhan, China. *General Hospital Psychiatry*, 32(3), 268–275.

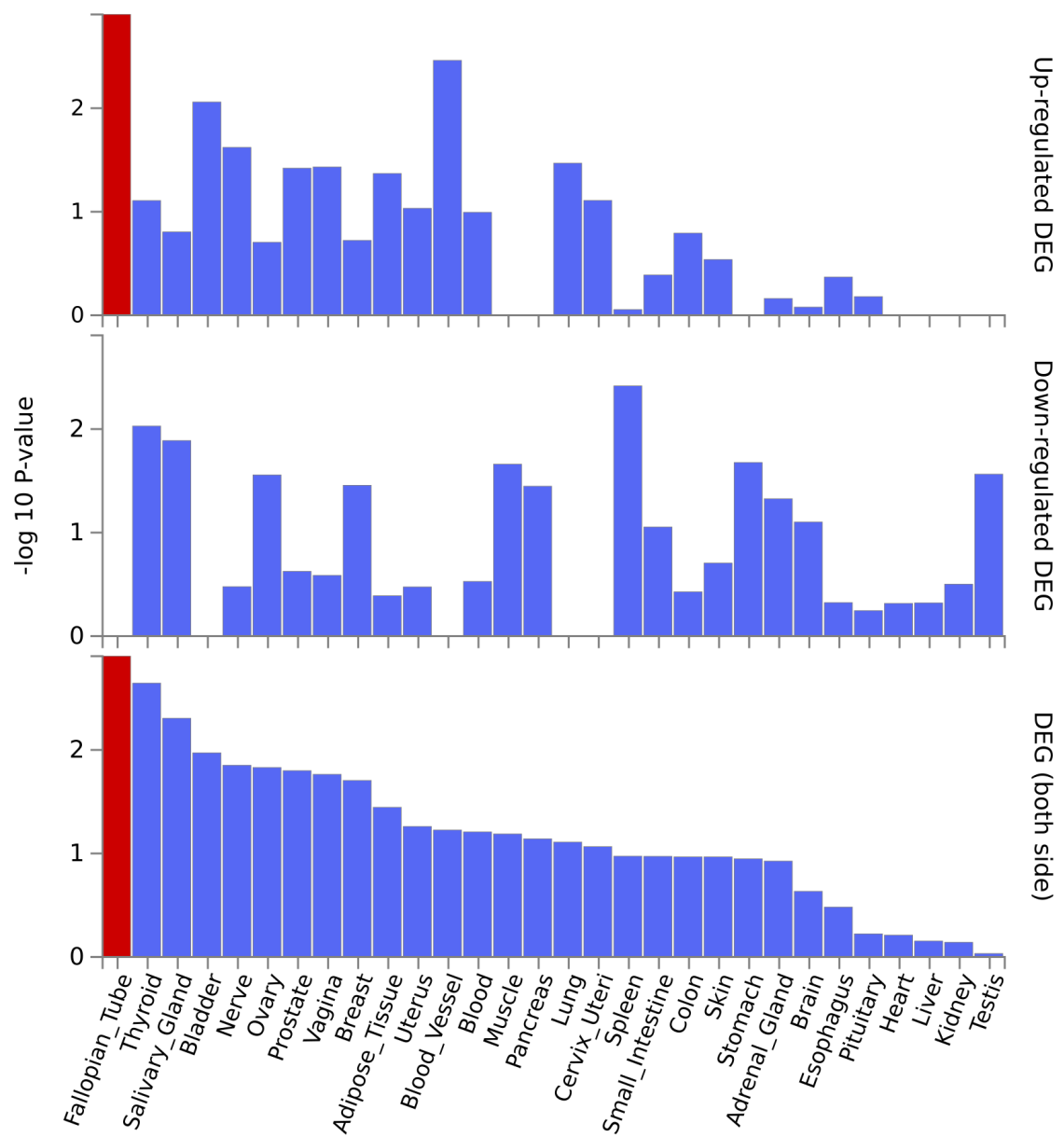
Zhou, X., Kang, L., Sun, X., Song, H., Mao, W., Huang, X., Zhang, Y., et al. (2013). Risk factors of mental illness among adult survivors after the Wenchuan earthquake. *Social Psychiatry and Psychiatric Epidemiology*, 48(6), 907–915.

Zis, P., Yfanti, P., Siatouni, A., Tavernarakis, A., & Gatzonis, S. (2014). Determinants of depression among patients with epilepsy in Athens, Greece. *Epilepsy & Behavior*, 33, 106–109.

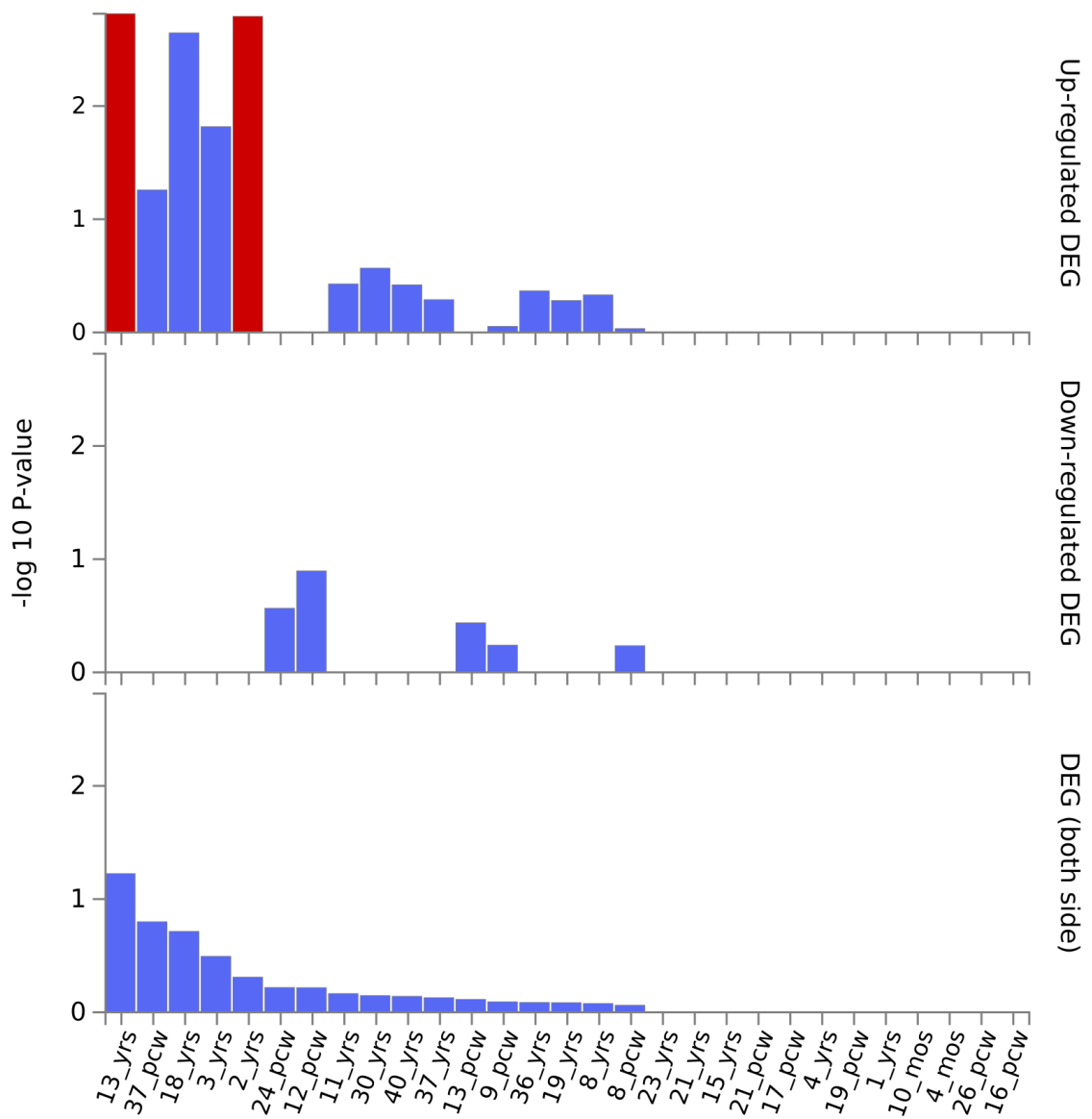
Appendix 1: FUMA analyses of SIMD EWAS DMP genes



Gene expression heatmap of genes associated with DMPs from the SIMD EWAS.

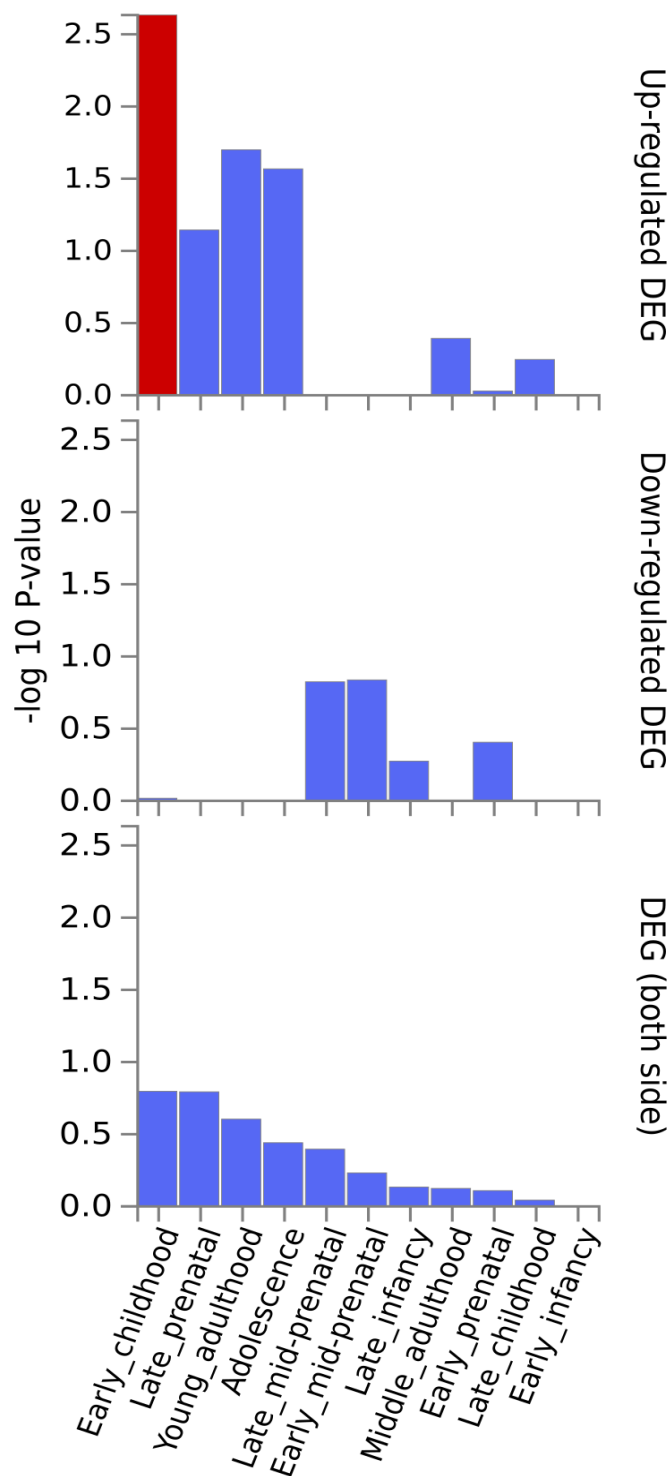


GTEx v6 30 general tissue types. Significantly enriched DEG sets ($P_{\text{bon}} < 0.05$) are highlighted in red.



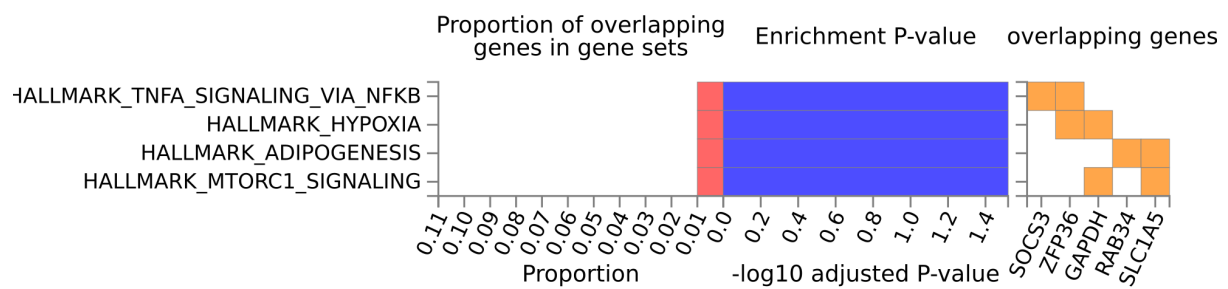
GTEx v8 30 general tissue types. Significantly enriched DEG sets ($P_{\text{bon}} < 0.05$) are highlighted in red.

5.5.1.1

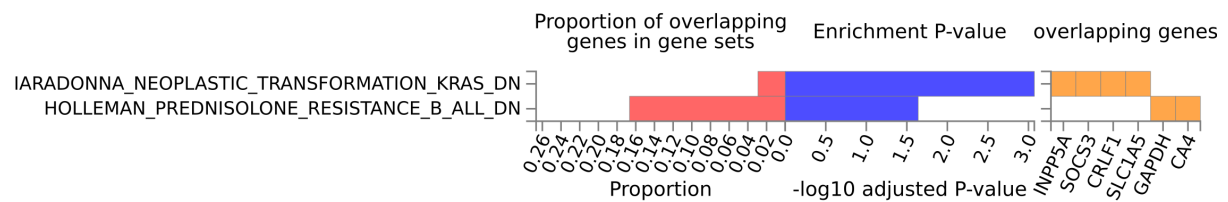


BrainSpan 11 general developmental stages of brain samples. Significantly enriched DEG sets ($P_{\text{bon}} < 0.05$) are highlighted in red.

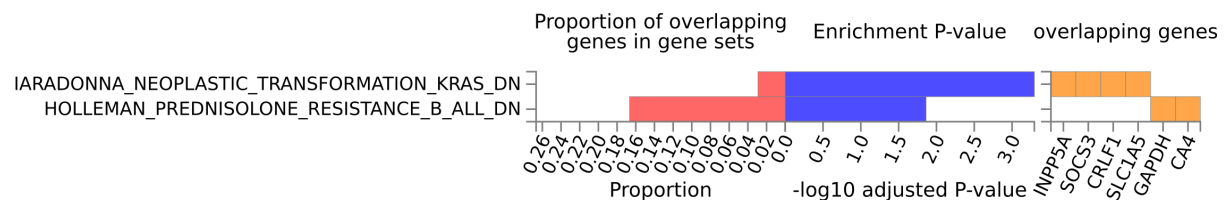
5.5.1.2



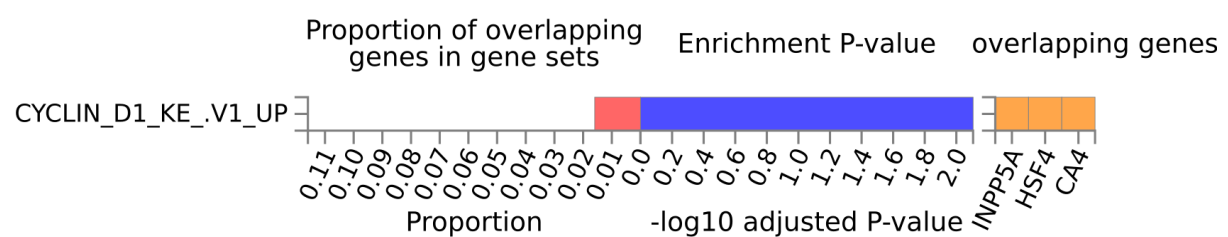
Hallmark gene sets (MsigDB h). Only gene sets with adjusted P-value < 0.05, are displayed.



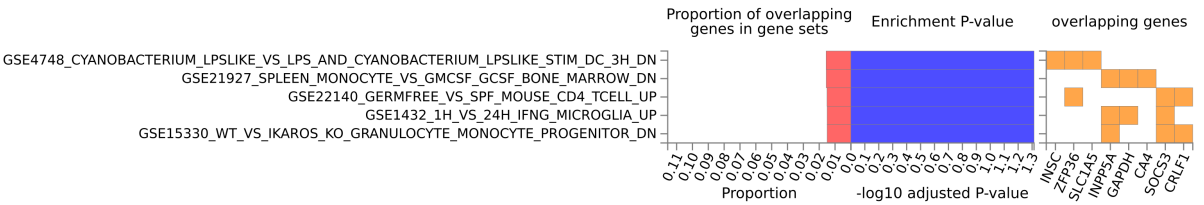
Curated_gene_sets. Only gene sets with adjusted P-value < 0.05, are displayed. Only gene sets with adjusted P-value < 0.05, are displayed.



Chemical and Genetic pertubation gene sets (MsigDB c2). Only gene sets with adjusted P-value < 0.05, are displayed.



Oncogenic signatures (MsigDB c6). Only gene sets with adjusted P-value < 0.05, are displayed.



Immunologic signatures (MsigDB c7). Only gene sets with adjusted P-value < 0.05, are displayed.

Appendix 2: FUMA analyses of SIMD EWAS DMR genes

List of DMRs from SIMD EWAS

region	chr	geneSymbol	dmr.start	dmr.end	dmr.z	dmr.p.adjust	estimate
58	17	ZNF232	5019056	5019989	-9.205947	2.77E-14	-4.06E-06
15	6	GNMT	42928337	42928773	-8.8018711	1.10E-12	-5.78E-06
91	12	GAPDH	6642201	6642953	8.7675888	1.49E-12	3.70E-06
27	16	ESRP2	68270220	68270882	-8.6991181	2.74E-12	-2.65E-06
93	19	SLC1A5	47287964	47288263	8.5131496	1.38E-11	3.78E-06
132	10	ADO	64565336	64565772	8.48871791	1.71E-11	3.46E-06
19	2		45159504	45160759	-8.4753516	1.92E-11	-3.12E-06
46	7	LOC650226	56515510	56516309	-8.3777302	4.41E-11	-6.10E-06
17	17	ACSF2	48545805	48546620	-8.3606017	5.10E-11	-5.53E-06
17	17	CHAD	48545805	48546620	-8.3606017	5.10E-11	-4.21E-06
199	17	SOCS3	76354621	76355202	8.30356445	8.26E-11	5.12E-06
98	1	G0S2	209848479	209848864	-8.2400541	1.41E-10	-8.00E-06
30	5	GPX3	150399909	150400531	-8.0732831	5.60E-10	-3.42E-06
7	4	MARCH1	165304473	165305815	-7.9912835	1.09E-09	-3.29E-06
62	22	HORMAD2	30476098	30476345	-7.9841868	1.16E-09	-2.72E-06
4	13	EDNRB	78492306	78493203	-7.9345841	1.73E-09	-3.37E-06
23	8	RGS22	101117952	101118899	-7.9324752	1.76E-09	-3.05E-06
10	17	RAB34	27044629	27045022	-7.8465267	3.50E-09	-4.58E-06
1	11	KCNQ1DN	2890473	2891098	-7.8304167	3.98E-09	-5.19E-06
3	6	TUBB	30689844	30691914	7.8117978	4.61E-09	1.27E-06
190	2	SPRED2	65593761	65594021	7.77606387	6.12E-09	3.83E-06
12	5	ZNF300	150284416	150285117	-7.7597519	6.96E-09	-2.57E-06
137	1	ARTN	44398998	44399428	-7.7203431	9.49E-09	-3.77E-06
37	1	FAAH	46859780	46860511	-7.6818593	1.28E-08	-4.38E-06
28	1	PEX10	2345398	2345475	7.64206208	1.75E-08	4.03E-06
16	16	DNASE1L2	2286370	2287763	-7.618443	2.10E-08	-1.97E-06
14	6	GTF2H4	30881464	30881766	-7.5630771	3.22E-08	-5.85E-06
14	6	VARs2	30881464	30881766	-7.5630771	3.22E-08	-2.62E-06
153	5	C5orf49	7851256	7851674	-7.5189741	4.52E-08	-3.10E-06

186	19	MYBPC2	50936058	50936172	-7.5064468	4.97E-08	-5.32E-06
330	19	TM6SF2	19383838	19384165	-7.4968544	5.35E-08	-3.87E-06
145	2		63285491	63286621	-7.4822453	5.98E-08	-3.08E-06
33	19	RSPH6A	46318439	46319398	-7.4051454	1.07E-07	-2.99E-06
33	19	SYMPK	46318439	46319398	-7.4051454	1.07E-07	-1.71E-06
56	13	SLITRK5	88328009	88329453	-7.4026629	1.09E-07	-4.77E-06
268	2		8683604	8684126	7.39624747	1.15E-07	4.72E-06
375	15		70740429	70740470	-7.3905597	1.20E-07	-5.57E-06
63	6	MDC1	30686652	30687511	7.38948526	1.21E-07	3.66E-06
63	6	TUBB	30686652	30687511	7.38948526	1.21E-07	2.39E-06
105	3		188665413	188665440	-7.3601403	1.50E-07	-3.56E-06
105	3	TPRG1-AS1	188665413	188665440	-7.3601403	1.50E-07	-3.64E-06
2	1	TP73	3567732	3568245	-7.3444997	1.69E-07	-3.46E-06
187	20	PAK7	9819272	9819805	-7.3014076	2.33E-07	-4.87E-06
114	7	AQP1	30951272	30951498	-7.2446254	3.55E-07	-3.06E-06
26	14		48144314	48145588	-7.2313149	3.91E-07	-3.75E-06
26	14	MDGA2	48144314	48145588	-7.2313149	3.91E-07	-3.60E-06
307	8		93522056	93522271	-7.2302955	3.94E-07	-4.16E-06
103	3	CCDC37	126113640	126113784	-7.2164263	4.37E-07	-3.77E-06
99	2	C2orf39	26624612	26624865	-7.2124215	4.50E-07	-3.27E-06
208	1	SGIP1	66999588	66999813	-7.1783095	5.77E-07	-4.43E-06
378	16	HSF4	67199463	67199830	-7.1590598	6.64E-07	-5.22E-06
141	1		156357839	156358119	-7.1496463	7.12E-07	-3.60E-06
45	6	KHDRBS2	62996022	62997123	-7.1251355	8.51E-07	-4.70E-06
13	6		29521272	29521430	-7.0989886	1.03E-06	-7.83E-06
95	1	DNALI1	38022327	38022657	-7.0971941	1.04E-06	-2.00E-06
368	11	GAL3ST3	65816642	65816819	-7.0938228	1.07E-06	-2.38E-06
386	19	APC2	1456157	1456337	-7.0902557	1.10E-06	-4.93E-06
144	1	ACTA1	229569608	229569902	-7.088739	1.11E-06	-5.36E-06
76	5	ZNF454	178367827	178368620	-7.0851657	1.14E-06	-2.87E-06
204	1	C1orf94	34642305	34642862	-7.0499971	1.46E-06	-3.80E-06
112	6	SLC26A8	35992082	35992691	-7.0435793	1.53E-06	-1.64E-06
59	17	FBXO39	6679264	6679499	-6.9986962	2.11E-06	-3.96E-06
205	1	EVA1B	36789768	36790060	-6.9804037	2.41E-06	-1.86E-06
205	1	FAM176B	36789768	36790060	-6.9804037	2.41E-06	-1.55E-06
271	7	AHCYL2	129007902	129008407	6.97787133	2.45E-06	2.89E-06
200	1	LOC102724312	1369793	1369950	-6.9672874	2.64E-06	-2.02E-06
200	1	VWA1	1369793	1369950	-6.9672874	2.64E-06	-3.88E-06
34	20	SYNDIG1	24449136	24449734	-6.9552099	2.88E-06	-7.53E-06

34	20	TMEM90B	24449136	24449734	-6.9552099	2.88E-06	-4.94E-06
156	5	CKMT2	80529100	80529340	-6.9156052	3.81E-06	-2.31E-06
156	5	RNU5E	80529100	80529340	-6.9156052	3.81E-06	-2.97E-06
259	17	RARA	38501412	38501677	-6.8773178	4.99E-06	-3.77E-06
83	12	SUOX	56390939	56391451	-6.8755611	5.05E-06	-3.30E-06
318	12	CAMKK2	121735363	121735631	-6.8734959	5.12E-06	-3.44E-06
115	10	DOCK1	128593624	128594417	-6.8627236	5.53E-06	-5.09E-06
66	1	MARCH1	220921306	220922217	-6.8615569	5.57E-06	-5.15E-06
66	1	MOSC2	220921306	220922217	-6.8615569	5.57E-06	-2.61E-06
288	2		203036093	203036235	-6.8503032	6.03E-06	-2.41E-06
87	17	DNAH9	11501632	11502254	-6.8480217	6.13E-06	-2.99E-06
388	19		57617602	57618136	-6.8329961	6.80E-06	-6.15E-06
43	5	TGFBI	135364552	135364970	-6.8213022	7.38E-06	-4.72E-06
35	6	C6orf48	31804078	31804883	6.82059975	7.42E-06	2.73E-06
70	3	GRM7	6902843	6903327	-6.8196914	7.46E-06	-2.62E-06
220	4	CXCL1	74734762	74735092	-6.8090117	8.04E-06	-4.01E-06
111	6	C6orf146	4078943	4079647	-6.7849754	9.50E-06	-4.19E-06
111	6	C6orf201	4078943	4079647	-6.7849754	9.50E-06	-3.36E-06
276	12	KRT80	52585844	52586127	6.78485581	9.51E-06	2.37E-06
358	5	ITGA2	52286121	52286162	-6.7613437	1.12E-05	-3.98E-06
68	2	LOC654342	91847976	91848218	-6.7436887	1.26E-05	-2.41E-06
20	2	GULP1	189156425	189157513	-6.7383069	1.31E-05	-2.58E-06
193	7	UPP1	48127997	48128242	6.73472495	1.34E-05	3.24E-06
75	5	SNCAIP	121646553	121647856	-6.7273728	1.41E-05	-4.00E-06
262	19	EPHX3	15344186	15344364	-6.7229038	1.46E-05	-3.09E-06
179	15	PYGO1	55881044	55881089	-6.7103778	1.59E-05	-4.80E-06
21	4		53617184	53618129	-6.7079893	1.61E-05	-2.70E-06
21	4	ERVMER34-1	53617184	53618129	-6.7079893	1.61E-05	-2.95E-06
127	19	HSPB6	36247526	36248077	-6.6886186	1.84E-05	-1.94E-06
127	19	PROSER3	36247526	36248077	-6.6886186	1.84E-05	-2.23E-06
44	5	LOC134466	150325954	150326862	-6.6860862	1.88E-05	-3.40E-06
9	13	SOX1	112720927	112723421	-6.6849141	1.89E-05	-2.16E-06
97	1	CR1L	207818368	207818449	-6.6755843	2.01E-05	-7.09E-06
18	3	ZNF385D	21792195	21793557	6.66048351	2.23E-05	3.56E-06
226	6	COL11A2	33160969	33161333	-6.6489334	2.41E-05	-3.82E-06
394	19	LDLR	11202610	11203039	6.61561424	3.03E-05	3.55E-06
131	10	KLF6	3823790	3824687	6.60327156	3.29E-05	1.81E-06
61	18	DSC3	28622474	28623060	-6.5934743	3.51E-05	-3.43E-06
72	5	MARCH1	16180055	16180393	-6.5772627	3.92E-05	-4.40E-06

192	7	NFE2L3	26192756	26193805	6.56272737	4.32E-05	4.22E-06
5	14		102025815	102026525	-6.5504866	4.69E-05	-2.11E-06
5	14	DIO3	102025815	102026525	-6.5504866	4.69E-05	-3.51E-06
67	2	POMC	25391656	25391830	-6.5333042	5.26E-05	-1.87E-06
78	7		63385990	63387147	-6.5295043	5.40E-05	-4.82E-06
206	1	ATPAF1	47134286	47134349	-6.5242766	5.59E-05	-5.07E-06
280	22	NCF4	37256947	37257594	6.51589723	5.91E-05	2.51E-06
180	16	A2BP1	6069009	6069731	-6.5136969	6.00E-05	-2.09E-06
180	16	RBFOX1	6069009	6069731	-6.5136969	6.00E-05	-3.56E-06
53	10	HHEX	94450826	94451736	-6.5079534	6.23E-05	-2.54E-06
251	15		53092788	53093707	-6.5024729	6.46E-05	-5.64E-06
94	19	RPL13A	49993758	49994239	6.49608455	6.74E-05	2.48E-06
94	19	RPL13AP5	49993758	49994239	6.49608455	6.74E-05	2.87E-06
94	19	SNORD34	49993758	49994239	6.49608455	6.74E-05	2.49E-06
94	19	SNORD35A	49993758	49994239	6.49608455	6.74E-05	2.59E-06
237	9	PALM2	112402639	112402944	-6.4939912	6.84E-05	-3.82E-06
256	17	RNF112	19314299	19314408	-6.4885273	7.09E-05	-4.17E-06
41	3	THRB	24536252	24537177	-6.4841962	7.30E-05	-3.46E-06
64	12	SLC6A12	322641	323499	6.47654714	7.68E-05	1.88E-06
188	21	IFNGR2	34774726	34775045	-6.4688788	8.08E-05	-1.73E-06
89	20	BMP7	55841149	55842071	-6.4517017	9.05E-05	-3.13E-06
367	11		32008961	32009163	-6.4495826	9.17E-05	-7.16E-06
29	2	FAM84A	14774086	14775187	-6.4409938	9.71E-05	-2.67E-06
29	2	LOC653602	14774086	14775187	-6.4409938	9.71E-05	-2.27E-06
396	10	EBF3	131697080	131697193	-6.4360748	0.00010029	-9.89E-06
133	1	CLSTN1	9884565	9884732	-6.4322153	0.00010287	-3.87E-06
108	4	PF4V1	74718978	74719597	-6.4204104	0.00011117	-4.71E-06
248	14		74100798	74100987	-6.4196993	0.00011169	-2.99E-06
106	4		298743	299311	-6.4066775	0.00012166	-2.64E-06
189	22	C22orf45	24891129	24891166	-6.403923	0.00012387	-7.76E-06
349	2		91873998	91874634	-6.4002364	0.0001269	-2.78E-06
370	11	ESAM	124629580	124629925	-6.3957351	0.0001307	-4.11E-06
379	16	DPEP2	68033859	68034227	-6.3868812	0.00013849	-2.37E-06
203	1	ESPNP	17018684	17019133	-6.3847421	0.00014044	-4.54E-06
130	3	RASSF1	50375624	50376409	6.38443379	0.00014072	2.62E-06
109	4		141419276	141419833	-6.3637802	0.00016102	-5.00E-06
109	4	MGAT4D	141419276	141419833	-6.3637802	0.00016102	-3.93E-06
325	17		27918207	27918580	-6.3615438	0.00016339	-3.37E-06
85	15	C15orf62	41061689	41062224	-6.3611021	0.00016386	-1.68E-06

85	15	DNAJC17	41061689	41062224	-6.3611021	0.00016386	-2.34E-06
121	16	CSDAP1	31580150	31581309	-6.3500652	0.00017605	-2.89E-06
40	2	GBX2	237076135	237077265	-6.3483592	0.00017802	-3.02E-06
283	1	TDRD5	179560983	179561347	-6.3447577	0.00018223	-7.73E-06
143	1	ESRRG	217310887	217311547	-6.344694	0.00018231	-2.49E-06
218	3	C3orf25	129147397	129147616	-6.3418456	0.00018571	-5.29E-06
123	17	RARA	38498098	38498912	-6.341127	0.00018658	-2.89E-06
123	17	RARA-AS1	38498098	38498912	-6.341127	0.00018658	-4.12E-06
183	17	VMO1	4689640	4689879	-6.3354	0.00019364	-4.82E-06
55	11		64739320	64739636	-6.3259289	0.0002059	-3.77E-06
55	11	C11orf85	64739320	64739636	-6.3259289	0.0002059	-3.56E-06
217	3	C3orf52	111804832	111805034	-6.3144617	0.00022176	-2.11E-06
320	12		125038763	125039343	-6.3052371	0.00023538	-2.14E-06
366	11	APBB1	6425978	6426039	-6.2980359	0.00024658	-4.69E-06
287	2		121411889	121412432	-6.296165	0.00024957	-4.44E-06
213	3	CPNE9	9745411	9746059	-6.2957924	0.00025017	-3.16E-06
207	1	DMRTA2	50886411	50886969	-6.2862435	0.00026605	-4.74E-06
92	17	RPL23A	27048223	27049557	6.271992	0.00029159	1.60E-06
92	17	SNORD4A	27048223	27049557	6.271992	0.00029159	2.42E-06
71	4	GLRB	157996959	157997750	-6.2686527	0.00029791	-3.24E-06
139	1		91301731	91301962	-6.2597742	0.00031537	-7.34E-06
390	4		693321	693333	6.25600267	0.00032309	6.31E-06
96	1	LPPR4	99729593	99730221	-6.255156	0.00032485	-4.23E-06
362	6		100917397	100917416	-6.2487787	0.00033839	-4.35E-06
198	17	LRRC59	48473505	48473941	6.24427786	0.00034828	2.28E-06
399	19	GGN	38877874	38877908	-6.2405274	0.00035673	-4.48E-06
225	5	ZFP2	178322752	178322924	-6.2393824	0.00035935	-3.47E-06
165	7		121945440	121946608	-6.2323135	0.00037595	-3.73E-06
165	7	FEZF1	121945440	121946608	-6.2323135	0.00037595	-3.64E-06
311	10	FGFR2	123357603	123357754	-6.2315355	0.00037783	-4.64E-06
73	5	ESM1	54281198	54281687	-6.2296663	0.00038236	-4.11E-06
315	12	IQSEC3	175809	176189	-6.2284185	0.00038542	-3.08E-06
51	8		69242921	69243904	-6.22517	0.00039349	-2.64E-06
51	8	C8orf34	69242921	69243904	-6.22517	0.00039349	-5.76E-06
51	8	C8orf34-AS1	69242921	69243904	-6.22517	0.00039349	-1.27E-06
201	1	RNF207	6268709	6269448	-6.2191109	0.00040899	-4.63E-06
172	11	DGKZ	46353999	46354302	-6.216892	0.00041481	-2.85E-06
255	17	ATP1B2	7554666	7555411	-6.2142769	0.00042178	-3.40E-06
113	6	RAB32	146864388	146865263	-6.2111549	0.00043024	-3.26E-06

196	13	MIR17HG	92001764	92003278	6.20987051	0.00043378	3.42E-06
196	13	MIR19A	92001764	92003278	6.20987051	0.00043378	4.60E-06
110	5	ADAMTS16	5139866	5140646	-6.2096002	0.00043452	-3.96E-06
164	7	WNT16	120968877	120969243	-6.2072769	0.00044099	-5.14E-06
331	19	SIGLEC10	51920885	51921017	-6.1910285	0.00048895	-2.54E-06
241	11	FEZ1	125365004	125365588	-6.1901587	0.00049165	-3.90E-06
90	4	MSX1	4864110	4864902	6.17781505	0.00053166	3.91E-06
155	5	ISL1	50678712	50679151	-6.1777953	0.00053173	-4.07E-06
36	20	BLCAP	36148738	36149271	6.16873748	0.00056309	2.04E-06
185	19	MATK	3801727	3801874	-6.1639624	0.00058035	-3.59E-06
184	17	CA10	50235393	50236104	-6.1588844	0.00059927	-3.93E-06
235	8	DPYS	105478683	105479318	-6.1563034	0.00060911	-4.17E-06
381	17	ZFP3	4981221	4981416	-6.1562874	0.00060917	-3.91E-06
214	3	PPARG	12328656	12328740	-6.1560389	0.00061013	-5.76E-06
270	7	GARS	30635712	30635889	6.15239369	0.00062432	5.94E-06
324	17	CENPV	16257111	16257165	-6.1502289	0.0006329	-3.93E-06
211	1	C1orf106	200860363	200860537	-6.1330539	0.00070519	-4.83E-06
281	1	MST1P9	17086061	17086071	-6.1184127	0.00077312	-4.49E-06
347	1		234905686	234906223	-6.1062568	0.00083432	-2.86E-06
223	5		101119084	101119766	-6.1053597	0.00083903	-8.30E-06
334	1	NTRK1	156785467	156785798	6.10525543	0.00083957	1.75E-06
334	1	SH2D2A	156785467	156785798	6.10525543	0.00083957	1.48E-06
374	15	MIR548H4	69373153	69373260	-6.1025089	0.00085413	-3.98E-06
352	2		240884646	240884925	-6.1011081	0.00086165	-4.08E-06
135	1	FABP3	31845904	31846110	-6.0941577	0.00089993	-2.27E-06
278	19	SLC1A5	47289503	47289611	6.09098148	0.00091797	3.18E-06
57	15	MIR1469	96875354	96876486	-6.0887066	0.00093111	-2.28E-06
57	15	NR2F2	96875354	96876486	-6.0887066	0.00093111	-2.25E-06
364	8	AGPAT5	6575798	6575881	-6.080981	0.0009771	-2.16E-06
356	4	IGFBP7	57975959	57976034	-6.0714635	0.00103682	-5.94E-06
197	16	MAZ	29820119	29820592	6.06950796	0.00104953	2.35E-06
169	8	EFCAB1	49647717	49648168	-6.0577431	0.00112921	-3.77E-06
239	10	PHYHIPL	60936070	60936925	-6.0573477	0.00113199	-2.23E-06
107	4	RHOH	40195003	40195318	-6.057301	0.00113232	-4.70E-06
228	6	PDE7B	136172755	136173085	-6.0555681	0.00114458	-2.47E-06
80	8	PENK	57358130	57358721	-6.0453355	0.00121964	-2.89E-06
246	13	SLC15A1	99404506	99405124	-6.0435889	0.00123293	-3.40E-06
171	10	C10orf26	104535792	104536121	-6.0426129	0.00124041	-9.53E-07
50	8	SOX17	55370192	55371829	-6.0376959	0.00127879	-3.26E-06

340	5	TRIP13	912712	912860	6.03721626	0.0012826	4.52E-06
176	14	ADCY4	24803925	24804342	-6.0271961	0.00136467	-3.34E-06
170	8	CCDC166	144790395	144790772	-6.0206299	0.0014212	-5.70E-06
170	8	LOC100130274	144790395	144790772	-6.0206299	0.0014212	-6.26E-06
272	8	SQLE	126011433	126012434	6.0189152	0.00143634	2.71E-06
253	16	B3GNT9	67184945	67185194	-6.0173478	0.00145031	-3.17E-06
219	4	OTOP1	4228239	4228650	-6.0167967	0.00145525	-2.09E-06
11	7	HOXA2	27142547	27143788	6.01647289	0.00145817	2.46E-06
118	12		133412002	133413532	-6.0125448	0.00149395	-3.37E-06
350	2	LOC440839	114048868	114048973	-6.0104356	0.00151352	-7.17E-06
77	6	GRIK2	101846767	101847656	-6.0055416	0.00155989	-2.60E-06
202	1	VPS13D	12538341	12538678	-6.0040066	0.00157472	-2.85E-06
159	5	PCDHGA4	140777446	140777611	-6.0027339	0.00158712	-3.94E-06
159	5	PCDHGB5	140777446	140777611	-6.0027339	0.00158712	-4.30E-06
245	13	LMO7	76334325	76334866	-5.9906382	0.00170979	-2.36E-06
39	2	LYPD1	133427612	133428523	-5.987247	0.00174581	-3.18E-06
138	1	AGBL4	50489418	50489827	-5.9826396	0.00179594	-2.58E-06
319	12	CDK2AP1	123753944	123754755	-5.9820434	0.00180252	-2.52E-06
54	10	NKX6-2	134600489	134600701	-5.9810615	0.00181342	-7.55E-06
102	3	PDZRN3	73673893	73674170	-5.9743059	0.00189018	-2.08E-06
313	11	DGKZ	46366643	46367033	-5.9665733	0.00198193	-1.71E-06
304	6	PGM3	83903927	83904091	-5.9556381	0.00211912	-3.26E-06
116	11	GRIA4	105480771	105481306	-5.9482018	0.00221765	-3.01E-06
81	8	ZFPM2	106330173	106331995	-5.9466936	0.00223817	-2.73E-06
32	13	SIAH3	46425712	46425999	-5.9443008	0.00227111	-3.71E-06
126	19	ZNF69	11998206	11999000	-5.9417554	0.00230666	-2.50E-06
323	17		5000803	5001281	-5.9412763	0.00231342	-1.30E-06
341	10	DUSP5	112260509	112261195	5.93834441	0.00235516	2.73E-06
154	5		50265407	50265924	-5.9369817	0.00237481	-3.48E-06
154	5	LOC100287592	50265407	50265924	-5.9369817	0.00237481	-3.25E-06
252	15	TM6SF1	83776769	83777082	-5.9275025	0.002516	-4.19E-06
74	5	EDIL3	83679941	83680719	-5.926626	0.00252946	-3.09E-06
238	10	NET1	5488436	5488628	-5.9146416	0.00272068	-2.26E-06
289	2	WNT10A	219744597	219744634	-5.9125963	0.00275469	-3.30E-06
48	7	WDR86	151107924	151108811	-5.8974984	0.00301889	-1.80E-06
339	4	TRIM2	154178264	154178415	5.8959279	0.00304774	6.30E-06
150	3	MME	154796736	154798159	-5.8954329	0.00305689	-2.29E-06
79	7	PPP1R9A	94536787	94537893	-5.8940212	0.00308314	-2.65E-06
275	11		43942418	43942604	5.89343077	0.00309418	2.04E-06

275	11	ALKBH3-AS1	43942418	43942604	5.89343077	0.00309418	1.71E-06
314	11		121593469	121593838	-5.8926407	0.00310902	-3.94E-06
117	12	ATP5G2	54071090	54071264	-5.8920453	0.00312024	-2.56E-06
124	18	CIDEA	12254148	12254556	-5.8892761	0.00317298	-3.79E-06
395	8	FAM135B	139509073	139509075	-5.8879242	0.00319903	-7.10E-06
142	1	LMX1A	165204745	165205003	-5.8878575	0.00320032	-4.57E-06
227	6		106433985	106434169	-5.8850205	0.00325571	-6.25E-06
24	12	EID3	104697532	104697983	-5.8844285	0.00326738	-4.73E-06
24	12	TXNRD1	104697532	104697983	-5.8844285	0.00326738	-6.40E-06
229	7		45001924	45002980	-5.882282	0.00331005	-1.85E-06
229	7	MYO1G	45001924	45002980	-5.882282	0.00331005	-4.73E-06
38	2	C2orf65	74875268	74876010	-5.8802157	0.00335164	-2.55E-06
38	2	M1AP	74875268	74876010	-5.8802157	0.00335164	-2.83E-06
286	2	LBX2	74726633	74726651	-5.8777838	0.00340124	-7.56E-06
60	18	PSMA8	23713650	23714084	-5.8766483	0.00342464	-6.33E-06
345	17	UNC13D	73840853	73840989	5.8745499	0.0034683	2.53E-06
6	14	DIO3	102026555	102027172	-5.8702836	0.00355874	-4.86E-06
152	4	CXCL6	74702175	74702781	-5.8615556	0.00375098	-3.01E-06
162	7	WBSCR17	70597091	70597921	-5.8597759	0.0037914	-2.52E-06
269	3	DALRD3	49057884	49058272	5.85378504	0.0039306	2.54E-06
269	3	NDUFAF3	49057884	49058272	5.85378504	0.0039306	3.95E-06
295	4	GALNTL6	172735549	172735924	-5.8456389	0.00412789	-4.23E-06
212	1	MARCH1	220959586	220960554	-5.839942	0.00427155	-5.37E-06
212	1	MOSC1	220959586	220960554	-5.839942	0.00427155	-2.96E-06
333	21	HLCS	38353114	38353273	-5.8284033	0.00457758	-4.27E-06
151	4	DRD5	9783196	9783246	-5.8211162	0.00478173	-3.90E-06
336	1		235116637	235116922	5.81693097	0.00490296	3.20E-06
69	2	ST6GAL2	107502677	107503819	-5.8120732	0.00504742	-3.41E-06
158	5	PCDHB4	140501277	140501661	-5.8101844	0.0051047	-5.86E-06
231	7	NOM1	156741514	156741808	-5.8096344	0.00512149	-5.45E-06
120	15		59157075	59157579	-5.801929	0.00536255	-2.87E-06
321	12	TMEM132C	128752040	128752573	-5.8002939	0.00541511	-4.68E-06
380	16	KIAA0182	85650374	85650478	-5.794459	0.00560676	-3.46E-06
279	20	ZNF217	52198164	52198378	5.78970049	0.00576793	2.38E-06
389	2		240866924	240867059	5.78704722	0.00585975	5.70E-06
8	10	LOC100128811	25464139	25465171	-5.7865724	0.00587633	-2.39E-06
119	13	ERICH6B	46189838	46190134	-5.7848071	0.00593837	-2.99E-06
119	13	FAM194B	46189838	46190134	-5.7848071	0.00593837	-4.11E-06
128	19	TMEM91	41882232	41882654	-5.7808698	0.00607905	-2.71E-06

263	19	PGLYRP2	15580310	15580700	-5.7744559	0.00631519	-3.86E-06
372	14	NIN	51290136	51290662	-5.7725117	0.00638851	-3.89E-06
309	10	C10orf107	63422630	63423079	-5.7630884	0.0067558	-4.36E-06
383	17		55888454	55888548	-5.7628015	0.0067673	-2.85E-06
52	9	ELAVL2	23821981	23822390	-5.7525153	0.00719235	-3.19E-06
261	18	CHST9	24765119	24765506	-5.7468881	0.00743575	-2.67E-06
312	11	MUC6	1036727	1036865	-5.742583	0.00762736	-7.43E-06
391	9	C9orf139	139924651	139924709	5.74224228	0.00764272	2.14E-06
148	3	CHL1	239229	239665	-5.7327607	0.00808264	-5.10E-06
306	8	NRG1	31496771	31497042	-5.7305672	0.00818787	-5.82E-06
125	18	NETO1	70534110	70535005	-5.7297378	0.008228	-2.96E-06
161	6	HTR1B	78173200	78174065	-5.7279769	0.00831385	-2.53E-06
243	12	CCDC62	123258441	123258876	-5.7274505	0.00833968	-2.09E-06
284	2	PXDN	1748675	1748970	-5.7258115	0.0084206	-3.57E-06
216	3	VGLL3	87039150	87040082	-5.7238117	0.00852038	-4.18E-06
326	17	RASL10B	34058527	34058597	-5.7226869	0.008577	-2.59E-06
166	7	LEP	127881143	127881344	-5.7196564	0.00873138	-3.37E-06
136	1	COL16A1	32169706	32170066	-5.7166349	0.00888799	-2.53E-06
301	5	NME5	137474830	137475177	-5.7158483	0.00892921	-2.24E-06
129	1	MUTYH	45806350	45806501	5.71278434	0.00909153	8.21E-06
65	1	CD34	208084030	208085455	-5.7068532	0.00941395	-2.17E-06
267	22	PARVG	44568423	44568770	-5.7040759	0.00956871	-1.74E-06
285	2	OTX1	63279645	63280376	-5.6992976	0.00984079	-1.96E-06
173	12	SLC2A14	8025582	8025807	-5.6958088	0.01004418	-5.75E-06
392	11	FAR1	13690122	13690194	5.69241028	0.01024623	6.98E-06
353	3		46853631	46853910	-5.6919031	0.01027672	-2.52E-06
342	10	INPP5A	134385407	134386090	5.68988378	0.01039898	2.36E-06
122	17	RARA	38465281	38465510	-5.6747247	0.01136309	-1.62E-06
300	5	PITX1	134367090	134367394	-5.6733894	0.01145206	-4.41E-06
236	9	CA9	35676056	35676365	-5.6601432	0.01237209	-2.58E-06
249	15	SNAP23	42787188	42787359	-5.6562426	0.01265643	-2.89E-06
303	6	DDR1	30851051	30851086	-5.6457499	0.01345317	-4.29E-06
101	2	B3GNT7	232260139	232260305	-5.6455009	0.01347266	-4.39E-06
305	7	TRIL	28995786	28995874	-5.6384663	0.01403466	-4.06E-06
290	3	ZNF660	44626304	44626697	-5.6338555	0.0144153	-3.13E-06
49	8	UNC5D	35092687	35094001	-5.6335896	0.01443756	-2.81E-06
369	11	POU2F3	120110688	120110887	-5.6309053	0.01466409	-3.21E-06
240	11	KCNA4	30038619	30038675	-5.6282169	0.01489444	-7.07E-06
167	7	CALD1	134464123	134464411	-5.6269936	0.01500041	-3.07E-06

338	3	SH3BP5	15311021	15311183	5.62306393	0.01534581	1.78E-06
157	5	FBN2	127873106	127873840	-5.6183462	0.01577068	-2.56E-06
274	11	SLC22A18AS	2920052	2920265	5.61820106	0.01578393	2.58E-06
230	7		56551007	56552034	-5.6155326	0.01602948	-4.49E-06
260	17	DHX40	57642721	57642763	-5.6151621	0.01606387	-5.05E-06
354	3	ALG1L2	129800391	129800578	-5.614923	0.01608609	-1.75E-06
297	5	CDH12	22853134	22853730	-5.6108903	0.01646553	-3.12E-06
140	1	RUSC1	155290535	155290813	-5.6098808	0.01656186	-2.18E-06
140	1	RUSC1-AS1	155290535	155290813	-5.6098808	0.01656186	-3.70E-06
234	8	GRHL2	102505556	102506635	-5.6075189	0.0167894	-4.23E-06
104	3	SOX14	137483063	137484220	-5.6057042	0.01696628	-1.86E-06
168	8		25905134	25906064	-5.6056877	0.01696789	-3.01E-06
382	17		46719577	46720050	-5.602366	0.01729639	-3.23E-06
221	4	UGT8	115518992	115519578	-5.6017347	0.01735952	-3.64E-06
47	7	CADPS2	122525949	122526843	-5.6004033	0.01749338	-3.51E-06
277	16	ITGAL	30485597	30485966	5.59893101	0.01764258	1.90E-06
322	15	APBA2	29407792	29408145	-5.5986338	0.01767285	-2.17E-06
100	2	ZSWIM2	187713964	187714279	-5.597272	0.01781217	-5.66E-06
348	1		239550283	239550578	-5.5887785	0.01870556	-6.65E-06
182	16	CDH8	62069317	62070184	-5.5879404	0.01879604	-2.46E-06
296	5	CDH18	19988544	19988800	-5.5857792	0.01903131	-6.03E-06
86	16	HS3ST2	22825343	22825740	-5.5848729	0.01913083	-3.28E-06
146	2	VWA3B	98703619	98703698	-5.5838632	0.01924228	-5.48E-06
160	6	HMGCLL1	55443833	55444091	-5.5803456	0.01963553	-2.93E-06
264	20	BTBD3	11871375	11871743	-5.5767408	0.02004661	-5.63E-06
371	13	SACS	24007545	24007560	-5.5724065	0.02055196	-2.59E-06
181	16		54323335	54324597	-5.5723998	0.02055275	-3.89E-06
361	6	RPP21	30313495	30313517	-5.5713876	0.02067254	-3.62E-06
233	8	RALYL	85094923	85095513	-5.5699512	0.0208437	-3.58E-06
363	7	CFTR	117119601	117119637	-5.5631699	0.02167046	-3.86E-06
355	3	ZIC4	147113918	147114406	-5.5580092	0.02232088	-3.16E-06
343	12		111840792	111841024	5.55677445	0.02247928	3.07E-06
147	2	HOXD8	176994365	176994665	-5.5518435	0.02312281	-4.72E-06
177	14	FOXA1	38063813	38064532	-5.5508997	0.02324802	-3.48E-06
210	1		157948536	157949084	-5.549669	0.02341225	-2.66E-06
317	12	PDE1B	54943700	54943915	-5.5466773	0.02381622	-1.74E-06
346	19		13951168	13951845	5.54529655	0.02400494	2.33E-06
308	9	C9orf150	12775640	12776218	-5.5399451	0.02475015	-3.17E-06
308	9	LURAP1L	12775640	12776218	-5.5399451	0.02475015	-3.29E-06

178	14	HHIPL1	100111263	100111354	-5.537494	0.02509893	-2.44E-06
291	3	ZIC4	147106010	147106890	-5.5358794	0.02533127	-3.22E-06
332	20	APCDD1L	57089953	57090236	-5.5348874	0.02547505	-2.30E-06
257	17	ACCN1	32484008	32484035	-5.533491	0.0256788	-2.88E-06
360	5	HK3	176326516	176326528	-5.5305549	0.02611237	-2.48E-06
250	15	MAP1A	43809486	43809865	-5.5297854	0.02622716	-2.16E-06
316	12	ITGB7	53591398	53591766	-5.5297339	0.02623488	-2.42E-06
351	2	CD302	160654086	160654236	-5.5279267	0.02650648	-4.14E-06
84	14		58332116	58332963	-5.5245717	0.02701798	-1.73E-06
84	14	SLC35F4	58332116	58332963	-5.5245717	0.02701798	-2.90E-06
293	4	UCHL1	41258794	41259283	-5.5233331	0.02720923	-3.27E-06
195	12	IQSEC3	248296	248889	5.52269501	0.02730826	7.89E-06
385	18	NOL4	31802484	31802545	-5.5224402	0.0273479	-3.85E-06
163	7		84815361	84816543	-5.5205062	0.02765064	-2.87E-06
359	5	HAND1	153857907	153858102	-5.5180723	0.02803625	-3.55E-06
328	18	TMEM200C	5890293	5891001	-5.5173221	0.02815614	-3.85E-06
254	16	MARVELD3	71659894	71660055	-5.5156262	0.02842904	-3.31E-06
298	5	GDNF	37839916	37840839	-5.5143338	0.02863872	-2.99E-06
298	5	GDNF-AS1	37839916	37840839	-5.5143338	0.02863872	-9.71E-06
337	3	BHLHE40	5023050	5023394	5.51386042	0.02871589	3.63E-06
149	3	SLC6A11	10857717	10858003	-5.5117608	0.02906064	-5.40E-06
282	1	LHX8	75602167	75602846	-5.5096012	0.02941943	-4.40E-06
25	12	EID3	104697126	104697514	-5.5080025	0.02968778	-4.46E-06
25	12	TXNRD1	104697126	104697514	-5.5080025	0.02968778	-4.44E-06
373	15	SHC4	49255257	49255455	-5.5032024	0.03050789	-3.87E-06
377	16	ERN2	23724428	23724774	-5.5031344	0.03051965	-2.77E-06
344	14	ASB2	94421976	94422013	5.50163026	0.03078123	2.77E-06
310	10		103538476	103539519	-5.497113	0.03157992	-3.14E-06
209	1	ADORA3	112058236	112058975	-5.4948219	0.03199266	-2.15E-06
209	1	TMIGD3	112058236	112058975	-5.4948219	0.03199266	-1.33E-06
174	12		22094371	22095330	-5.4898733	0.03290206	-3.08E-06
302	5	PCDHB9	140568340	140568856	-5.4887333	0.03311508	-2.07E-06
265	20	SSTR4	23015906	23015936	-5.4833328	0.03414253	-4.52E-06
398	18	ZNF532	56531132	56531159	-5.4826977	0.03426537	-3.96E-06
194	10	DNAJC9	75005362	75005991	5.48222191	0.03435768	4.14E-06
294	4	NKX6-1	85417910	85418584	-5.4792909	0.03493167	-2.44E-06
244	12		123636175	123636548	-5.4788288	0.03502299	-2.39E-06
365	10	CPXM2	125651085	125651391	-5.4746564	0.03585827	-5.55E-06
273	9	RAPGEF1	134585306	134585801	5.47424693	0.03594127	2.39E-06

222	5	THBS4	79330959	79331477	-5.4735646	0.03607999	-3.60E-06
384	18		20839953	20840464	-5.4734907	0.03609505	-3.67E-06
384	18	CABLES1	20839953	20840464	-5.4734907	0.03609505	-3.46E-06
31	11	LAYN	111410878	111411216	-5.4731452	0.03616552	-2.88E-06
22	6	EYA4	133561649	133562479	-5.4701518	0.0367817	-2.25E-06
191	6	PSMB8	32809960	32810706	5.46973752	0.03686778	2.07E-06
88	19	USP29	57630153	57630710	-5.4672307	0.03739281	-5.52E-06
247	14	GALNTL1	69726197	69726319	-5.4662078	0.03760913	-3.27E-06
376	15	FES	91427605	91427738	-5.4651789	0.03782793	-4.25E-06
215	3	UBE2E1	23851386	23851909	-5.4625898	0.03838399	-1.93E-06
397	16	HSD11B2	67465461	67465561	-5.4625409	0.03839457	-3.21E-06
134	1	GRHL3	24646022	24646205	-5.4601899	0.03890651	-3.25E-06
258	17	GAS2L2	34079665	34079998	-5.4586305	0.0392497	-1.62E-06
327	17	CD79B	62009628	62009835	-5.4561634	0.03979867	-3.11E-06
329	19	EMR3	14785593	14785776	-5.4530328	0.040506	-2.11E-06
224	5	MEGF10	126626362	126626557	-5.4487734	0.041488	-4.19E-06
175	12	DUSP6	89744471	89744701	-5.4483673	0.04158282	-2.68E-06
242	12	ANO2	5936877	5937105	-5.4479523	0.04167992	-3.78E-06
292	3	FGF12	192232468	192232863	-5.446607	0.04199626	-3.31E-06
335	1	GAS5	173835989	173836636	5.44047435	0.04346798	3.23E-06
335	1	SNORD75	173835989	173836636	5.44047435	0.04346798	3.47E-06
335	1	SNORD78	173835989	173836636	5.44047435	0.04346798	2.26E-06
42	4	SFRP2	154711171	154711675	-5.4402154	0.04353121	-1.89E-06
266	20	GATA5	61051021	61051039	-5.4392674	0.04376346	-5.28E-06
299	5	F2RL1	76115784	76115974	-5.4371852	0.04427779	-2.59E-06
82	11	ASCL2	2291264	2291754	-5.4345082	0.04494767	-5.87E-06
232	8	PREX2	68864523	68864782	-5.4330678	0.04531216	-4.30E-06
393	12	CHST11	104853100	104853274	5.42743867	0.04676423	3.20E-06
357	4	TBC1D9	141678125	141678312	-5.4241288	0.047639	-3.90E-06
387	19		23253834	23254381	-5.421676	0.04829745	-5.09E-06

Appendix 3: SIMD LASSO coefficients

coef.name	coef.value
(Intercept)	10.9408481
cg20210689	39.1987004
cg18002909	-43.2689677
cg24404533	-1.0131020
cg03536474	-30.0144013
cg02036261	-5.4213249
cg03510159	22.0078972
cg23112821	13.9646990
cg25902187	-1.0735118
cg12226028	-51.2794181
cg10813070	-1.7001592
cg04802657	-19.5601955
cg22320365	-2.7462062
cg05180443	-0.8372191
cg04258457	9.4543701
cg18181703	2.7067940
cg13380624	-13.1259366
cg19715410	-12.7623114
cg26666107	-39.8574341
cg11853788	-92.1943770
cg26230651	-33.9982012
cg14762973	-52.0387270
cg07835424	-26.9211675
cg16415870	-36.5407281
cg27589809	18.9407525
cg18104012	-0.8107232
cg03459543	139.3084074
cg10666406	24.4518149
cg11667258	-20.9558223
cg14151259	-99.7589235
cg06610484	-10.6808809
cg02171435	-52.7068773