Author:
Khouja, Jasmine N

Title:
Young adult e-cigarette use

who uses e-cigarettes, why, and what are the potential consequences?

Author:
Khouja, Jasmine N

Title:
Young adult e-cigarette use

who uses e-cigarettes, why, and what are the potential consequences?

Young adult e-cigarette use:

who uses e-cigarettes, why, and what are the potential consequences?

Jasmine Natalie Khouja

October 2020

School of Psychological Science

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy (PhD) in the Faculty of Life Sciences.

Word count: 50437

Abstract

E-cigarettes are popular and effective smoking cessation aids, but there are concerns about the consequences of their use. Specifically, young people who have never smoked before may use e-cigarettes and this may lead to smoking initiation. Additionally, there are fears that long-term e-cigarette use may cause health issues.

To fully understand the potential impact of e-cigarette use and to inform policy, it is vital to understand who uses e-cigarettes, why, and what the potential consequences of use are – questions which are explored in this thesis. To investigate the likely harms of e-cigarette use, I have used novel methods to explore whether e-cigarettes may act as a gateway to smoking and explore what the health consequences of e-cigarette use may be.

First, I conducted a meta-analysis of studies exploring whether e-cigarette use acts as a gateway to smoking among young people. Second, I explored which young adults are most likely to vape and whether different reasons for vaping are associated with continued smoking. Third, I explored whether there is a shared genetic liability for smoking initiation and e-cigarette use in a cohort of young adults. Fourth, I investigated whether nicotine use without exposure to cigarette smoke (e.g., via e-cigarettes) may cause smoking-related poor health outcomes using a variety of Mendelian randomisation methods in a cohort of adults. Finally, I began to explore the relationship between nicotine and BMI in an experimental study of adults.

Overall, I found little evidence that e-cigarettes act as a gateway to smoking, however, smokers and vapers share common characteristics and may share a genetic predisposition to risk-taking. There was little evidence of nicotine use causing poor health outcomes without tobacco smoke exposure. The evidence suggests policies should encourage smokers to switch to vaping. Further research is needed to explore these findings using emerging methods and data.

Dedication and Acknowledgements

I would like to thank my supervisors, Marcus Munafò, Amy Taylor and Robyn Wootton for their support throughout my PhD. Their assistance has been invaluable, and I am grateful for all the opportunities I have had as a result of their continued support.

I would also like to thank the Medical Research Council's Integrative Epidemiology Unit at the University of Bristol for funding my work. Additionally, I would like to thank Bristol Medical School's Department of Population Health Sciences and the School of Psychological Science for providing high quality training throughout my PhD which has allowed me to include such varied methods in my thesis.

The University of Bristol's Tobacco and Alcohol Research Group (TARG) has also guided, supported, and inspired me throughout my PhD and I cannot thank the members of TARG enough. I would particularly like to thank the co-directors (Marcus Munafò, Angela Attwood and Olivia Maynard) who encouraged me to apply for this studentship. I'd also like to thank my workstream leads, Amy Taylor, Robyn Wootton and Hannah Sallis, who were always available to share their knowledge and boost my confidence.

I am also grateful to my many office mates, friends and fellow PhD students over the past four years. They have each supported me in their own unique way. I would like to thank my housemate for encouraging me to continue when the work seemed too daunting.

Finally, I would like to thank my family. Without their financial and emotional support, I would not have reached this stage in my career. I dedicate this thesis to them.

Author’s Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: [redacted] DATE: 16.10.2020

Publications, Pre-Prints and Pre-Registered Study Protocols

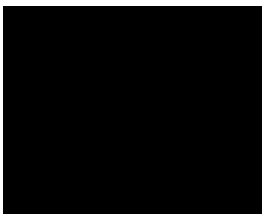
1a. Khouja, J. N., Suddell, S., Peters, S. E., Taylor, A., & Munafò, M. R. (2019). Does e-cigarette use in non-smoking young adults act as a gateway to smoking? A systematic review and meta-analysis. *OSF*; doi:10.17605/OSF.IO/3GC2Y.

1b. Khouja, J. N., Suddell, S. F., Peters, S. E., Taylor, A. E., & Munafò, M. R. (2020). Is e-cigarette use in non-smoking young adults associated with later smoking? A systematic review and meta-analysis. *medRxiv***;** doi:10.1101/19007005 %J**.**

1c. Khouja, J. N., Suddell, S. F., Peters, S. E., Taylor, A. E., & Munafò, M. R. (2020). Is e-cigarette use in non-smoking young adults associated with later smoking? A systematic review and meta-analysis. *Tobacco Control;* doi:10.1136/tobaccocontrol-2019-055433.

I took the lead in designing, implementing and interpreting the systematic review and meta-analysis with the support of my co-authors. Specifically, MM and AT provided feedback on the design of the study. Co-authors SS and SP assisted me with study selection and data extraction by independently checking 100% of my selection and extraction (50% each). I was responsible for the data analysis and was responsible for writing the first draft of the manuscript and editing the manuscript in response to comments from co-authors and reviewers. All co-authors provided feedback during the manuscript preparation.

Signed by first author: 

Signed by last author: 

2a. Khouja, J. N., Taylor, A. E., & Munafò, M. R. (2020). Associations between reasons for vaping and current vaping and smoking status: Evidence from a UK based cohort. *MedRxiv***;** doi:10.1101/19006007**.**

2b. Khouja, J. N., Taylor, A. E., & Munafò, M. R. (2021). Associations between reasons for vaping and current vaping and smoking status: Evidence from a UK based cohort. *Drug and Alcohol Dependence, 217,* 108362; doi.org/10.1016/j.drugalcdep.2020.108362.

I took the lead in designing, implementing and interpreting this observational study with the support of my co-authors. Specifically, MM and AT provided feedback on the design of the study and during the preparation of the manuscript. I was responsible for the data analysis and was responsible for writing the first draft of the manuscript and editing the manuscript in response to comments from co-authors and reviewers.

Signed by first author: 

Signed by last author:

3a. Khouja, J. N., Wootton, R. E., Taylor, A. E., Davey Smith, G., & Munafò, M. R. (2020). Association of genetic liability to smoking initiation with e-cigarette use in young adults. *MedRxiv*; doi: 10.1101/2020.06.10.20127464.

3b. Khouja, J. N., Wootton, R. E., Taylor, A. E., Davey Smith, G., & Munafò, M. R. (in press). Association of genetic liability to smoking initiation with e-cigarette use in young adults. *PLOS Medicine*.

I took the lead in designing, implementing and interpreting this study with the support of my co-authors. Specifically, MM, RW, AT and GDS provided feedback on the design of the study and during the preparation of the manuscript. I was responsible for the data analysis and was responsible for writing the first draft of the manuscript and editing the manuscript in response to comments from co-authors and reviewers.

Signed by first author: 

Signed by last author:

4. Khouja, J. N., Sanderson, E., Wootton, R. E., Taylor, A. E., & Munafò, M. R. (2021). A multivariable Mendelian randomisation study exploring the direct effects of nicotine on health compared with the other constituents of tobacco smoke: Implications for e-cigarette use. MedRxiv; doi: https://doi.org/10.1101/2021.01.12.21249493.

I took the lead in designing, implementing and interpreting this study with the support of my co-authors. Specifically, MM, RW, AT and ES provided feedback on the design of the study and during the preparation of the manuscript. I was responsible for the data analysis and was responsible for writing the first draft of the manuscript and editing the manuscript in response to comments from co-authors and reviewers.

Signed by first author: 

Signed by last author: 

5. Ferrar, J., **Khouja, J. N.,** Birch, L., Hamilton-Shield, J., Ness, A., Munafò, M. R., & Attwood, A. S. (2019). Effects of nicotine challenge on eating topography in non-dependent smokers. *OSF*; osf.io/nathz.

I assisted in designing and implementing this experimental study with the support of my co-authors and I took the lead in managing the study and in the analysis and interpretation of the results. Specifically, JF, LB, JHS AN, MM and AA provided feedback on the design of the study.

Signed by first author: 

Signed by last author: 

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Chapter Overview

Over the past decade, e-cigarettes have become extremely popular; in the UK alone there are currently an estimated 3.6 million users (Action on Smoking and Health, 2019a). Alongside this growing popularity, concerns about the safety and unintended consequences of use (particularly among young people) have also grown (Besaratinia & Tommasi, 2019). Despite evidence indicating that e-cigarettes are less harmful than cigarettes (Public Health England, 2015) and evidence suggesting they can aid smoking cessation (Hajek et al., 2019; Hartmann-Boyce et al., 2020), fears have arisen regarding the use of e-cigarettes by young people who have never smoked before.

E-cigarette users who have never smoked before appear to be at increased likelihood of subsequently smoking at least once in the future (Soneji, Barrington-Trimis, Wills, Leventhal, et al., 2017). However, it is unclear whether this association is causal. Furthermore, the extent to which these individuals are at greater risk is unclear due to heterogeneity between study effect estimates, low sample sizes, measurement limitations and inadequate control for potential confounders within the available studies (Glasser, Abudayyeh, Cantrell, & Niaura, 2018). It is important to understand which individuals use e-cigarettes, why, and how this may impact on behaviour in order to assess the potential consequences of use, such as smoking initiation, smoking continuation, or poor health outcomes. This understanding is necessary to determine whether cautious policy approaches (e.g., bans) are appropriate and could also help to inform public health strategies.

Although e-cigarettes contain far fewer, and lower levels of, the harmful toxicants and chemicals found in cigarettes (Goniewicz et al., 2014), the long-term health consequences of using e-cigarettes are currently unknown. As e-cigarettes are a relatively new product, it is impossible to observe the long-term effects of e-cigarette use. Instead, innovative methods are needed to explore the long-term effects without waiting decades for health outcome data to become available. Understanding the potential health effects of long-term e-cigarette use could help guide public health messages aimed at smoking cessation, given that a growing number of smokers believe

that e-cigarettes are equally or more harmful than cigarettes (Action on Smoking and Health, 2019a; Gravely et al., 2020).

The aim of this thesis is to understand which individuals are more likely to use e-cigarettes, why they use e-cigarettes, and what are the potential implications of e-cigarette use both in terms of continued use, the effect on smoking and health effects? I apply novel methods in the context of e-cigarette research – including using multivariable Mendelian randomisation and negative controls – to address these aims. Where possible, I focus on the association and effects among young adults (up to and including 30 years of age).

1.2 E-cigarettes

1.2.1 E-cigarette Terminology and Basic Function

E-cigarettes are known by a variety of names (e.g., electronic cigarettes, e-cigs, vapes, vape pens, vaporisers, mods, pods, JUULs, Electronic Nicotine Delivery Systems [ENDS], Alternative Nicotine Delivery Systems [ANDS]). The use of e-cigarettes can also be referred to as vaping or Juuling (among other terms). E-cigarettes are electronic devices which were introduced onto the UK market in 2007 (House of Commons Science and Technology Committee, 2018). They operate by heating a solution, often referred to as e-liquid or e-juice, which creates a vapour that can be inhaled via a mouthpiece. For the purposes of this thesis, I refer to these products as e-cigarettes, describe the use of these products as 'e-cigarette use' or 'vaping', and I refer to the solution as e-liquid.

1.2.2 Constituents of E-cigarette Vapour

E-cigarettes were originally marketed as a smoking cessation device because e-cigarettes can contain nicotine (the addictive component of cigarettes), but contain far fewer chemicals and toxicants (Goniewicz et al., 2014; Goniewicz et al., 2018). The chemicals and toxicants that are included in e-cigarettes are generally found at much lower levels than are found in cigarettes (Goniewicz et al., 2014). E-cigarettes are filled with e-liquid, which usually contains nicotine, flavours, and vegetable glycerin or propylene glycol (or a mixture of both), however, nicotine-free and unflavoured options are available. Nicotine can be in the form of a freebase solution or in the form of salts; nicotine is naturally found in tobacco as a salt and is deprotonated by adding an alkaline

chemical substance, such as ammonia, to free the nicotine base and create a purer form of nicotine – freebase nicotine (Gholap, Heyder, Kosmider, & Halquist, 2020). In e-cigarettes, nicotine salts can deliver high concentrations of nicotine to the user with less irritation than a freebase solution (Bowen & Xing, 2015). The flavourings included in e-liquids vary between products, and consequently the extent to which the resulting e-cigarette vapours are potentially harmful also varies; although the flavours are safe for oral consumption, there is often little evidence on the safety of these ingredients when heated and inhaled. In particular, cinnamon flavouring (created by adding cinnamaldehyde) has been shown to act as a respiratory irritant and is toxic when inhaled (Behar et al., 2016).

1.2.3 Generations of E-cigarettes

Over the past decade or so, e-cigarettes have developed rapidly in order to adapt to the needs of the consumer, resulting in four generations of e-cigarette (Figure 1.1). The first generation of e-cigarettes (also known as “cigalikes”) are similar in appearance to cigarettes but, for many users, deliver insufficient nicotine compared to cigarettes (Farsalinos et al., 2014). These devices can be disposable or rechargeable and can be refilled with prefilled cartridges. The second generation of e-cigarettes are pen-style – they deliver nicotine slightly better than first generation e-cigarettes. These devices are rechargeable and can be refilled with e-liquid. The third generation include ‘mods’ (modifiable devices), tank-style devices, and rebuildable dripping atomisers which allow users to modify their devices (e.g., by modifying the voltage) to suit their needs. These devices are also rechargeable and can be refilled with e-liquid. The latest devices are known as ‘pods’; pods are small (sometimes similar in appearance to a USB) but powerful devices that deliver nicotine as effectively as cigarettes using salt-based nicotine rather than freebase nicotine (Bowen & Xing, 2015). Nicotine delivery can also differ within each generation of e-cigarettes due to a range of factors, including the concentration of nicotine in the e-liquids used, atomiser resistance, wattage and user behaviour (e.g., the frequency of use, depth of inhalations etc.) (DeVito & Krishnan-Sarin, 2018).

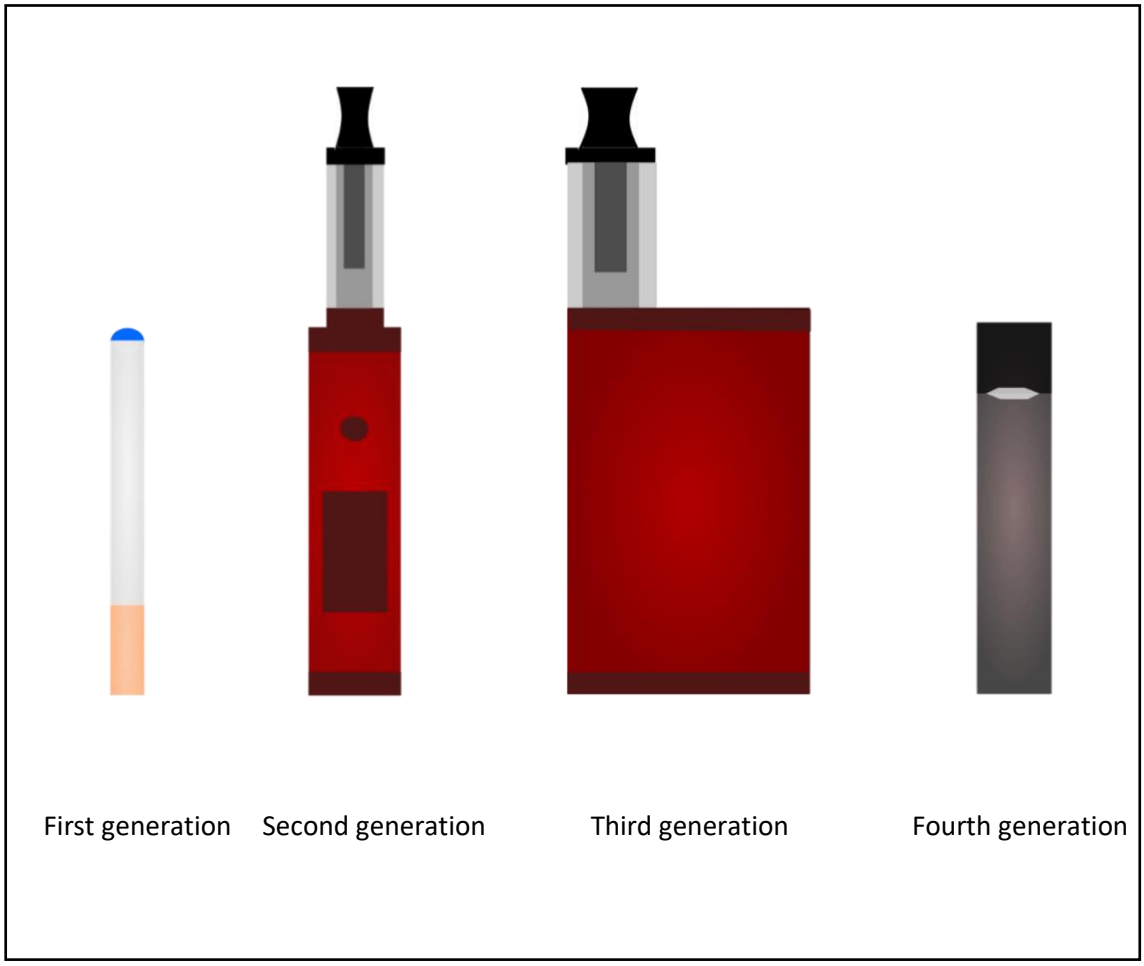


Figure 1.1. Appearance of varying e-cigarette types.

1.2.4 E-cigarette Policy

Due to safety concerns, e-cigarettes have been banned in a number of countries, including Australia and India (Chakma, Kumar, Bhargava, & Khanna, 2020; Gartner & Bromberg, 2019). In countries where e-cigarettes are not banned (e.g., the UK), there are often restrictions placed on the design and marketing of products. In the UK, e-cigarettes are currently regulated under the European Union Tobacco Products Directive (TPD). However, with the UK's exit from the European Union and the transition period ending soon, these regulations could be subject to change in the near future. At present, e-cigarettes have to be registered prior to sale (with a full list of ingredients and safety profile), e-liquid bottles must be child-proof and not exceed 10 ml, and nicotine concentrations cannot exceed 20 mg/ml of nicotine (Tobacco Products Directive, 2014). These regulations are stricter than in the US. Many countries have age restrictions on purchasing e-cigarettes – in the UK the minimum age of sale is 18 years of age, and federal law in the US prohibits the sale of tobacco products (including e-cigarettes) to those under 21 years of age.

Although e-cigarettes were originally marketed as a smoking cessation device, many countries now prohibit or restrict the advertising of e-cigarettes as cessation products. For example, the UK have placed restrictions on which devices can claim they are smoking cessation aids; only devices which are licensed as medicines by the Medicines and Healthcare products Regulatory Agency (MHRA) can make claims about smoking cessation. Currently, there are no available devices which are medically licensed.

1.3 E-cigarette Use and Smoking

1.3.1 Smoking Cessation

In a meta-analysis which aimed to evaluate the safety and effectiveness of using e-cigarettes to quit smoking (Hartmann-Boyce et al., 2020), there was moderate-certainty evidence to suggest that e-cigarettes can aid long-term smoking cessation (more than 6 months) compared with nicotine replacement therapy (N = 1,498 participants, risk ratio = 1.69, 95% CI 1.25 to 2.27) by Grading of Recommendations, Assessment, Development and Evaluation (GRADE) standards

(https://gdt.gradepro.org/app/handbook/handbook.html). This equates to an additional four successful quitters per 100 attempts.

Of the three randomised controlled trials included in the meta-analysis (all of which were low risk of bias), only one was conducted in the UK. In a randomised controlled trial of 886 attendees of UK NHS stop smoking services, Hajek and colleagues (2019) found that when combined with behavioural support, e-cigarettes were more effective for smoking cessation than nicotine replacement therapy. Of those randomly assigned to the e-cigarette arm of the study (who received an e-cigarette starter kit), 18% were abstinent at 1 year follow-up, whereas only 9.9% of those assigned to the nicotine replacement therapy group (who received a 3-month supply of nicotine replacement therapy as per standard treatment) were abstinent at 1 year follow-up.

A partially double-blind randomised controlled trial comparing nicotine containing e-cigarette use, nicotine-free e-cigarette use or no e-cigarette use (all in combination with individual counselling) is also in progress in Canada, and was due to complete data collection (with a target sample of 376 smokers who are willing to quit) in September 2020 (Hébert-Losier, Filion, Windle, & Eisenberg, 2020). Double-blinding was only possible for treatment allocation in the e-cigarette groups. Another double-blind randomised controlled trial comparing nicotine containing e-cigarettes to e-cigarettes without nicotine, and varenicline – a prescription drug which reduces cravings for nicotine and blocks the rewarding and reinforcing effects of smoking (West, Baker, Cappelleri, & Bushmakin, 2008) – is being conducted in France with 635 smokers who are willing to quit (Berlin et al., 2019). The growing evidence base should provide a clearer indication of the extent to which e-cigarettes are effective as a smoking cessation aid, but the current evidence suggests that nicotine containing e-cigarettes are an effective smoking cessation tool.

1.3.2 Observed Longitudinal Associations between Smoking and Vaping

With the potential e-cigarettes have to act as a smoking cessation tool, it would be logical to assume there is a strong association between smoking and later vaping. As expected, East and colleagues (2018) found that ever smoking young people in the UK are 3.54 times more likely to subsequently vape than never smokers. Biochemically verified adolescent smokers have also been found to be 7.24 times more likely to

subsequently vape in young adulthood (Khouja, Munafò, et al., 2020). However, there also appears to be a strong association between vaping and later smoking; young people who were ever vapers but never smokers were found to be approximately 12 times more likely to subsequently smoke (East et al., 2018).

A meta-analysis exploring the association between vaping among never smokers and subsequent smoking also found a strong positive association; adolescents and young adults were 3.5 times more likely to subsequently smoke (Soneji, Barrington-Trimis, Wills, Leventhal, et al., 2017). This suggests that there may be a causal influence of e-cigarette use on later smoking among never smokers. However, in their review of the impact of e-cigarette use on smoking, Glasser and colleagues (2018) highlighted that many of the studies included in the meta-analysis were at risk of bias, and there was also moderate heterogeneity between study estimates. Many of the included studies were based on a small number of e-cigarette users (Primack, Soneji, Stoolmiller, Fine, & Sargent, 2015) and used inappropriate methods (i.e., assessed experimental smoking rather than established smoking and did not adequately control for potential confounding). Therefore, it is misleading to draw any clear conclusions regarding causality from the findings of Soneji and colleagues (2017). There has been considerable discussion of potential explanations for the observed associations between e-cigarette use and smoking; for example, it has been proposed that vaping may normalise smoking, act as a gateway to smoking, or share a common liability with smoking (Chapman, Bareham, & Maziak, 2019; Etter, 2018; Sæbø & Scheffels, 2017). Each of these are explored in detail below and shown in Figure 1.2.

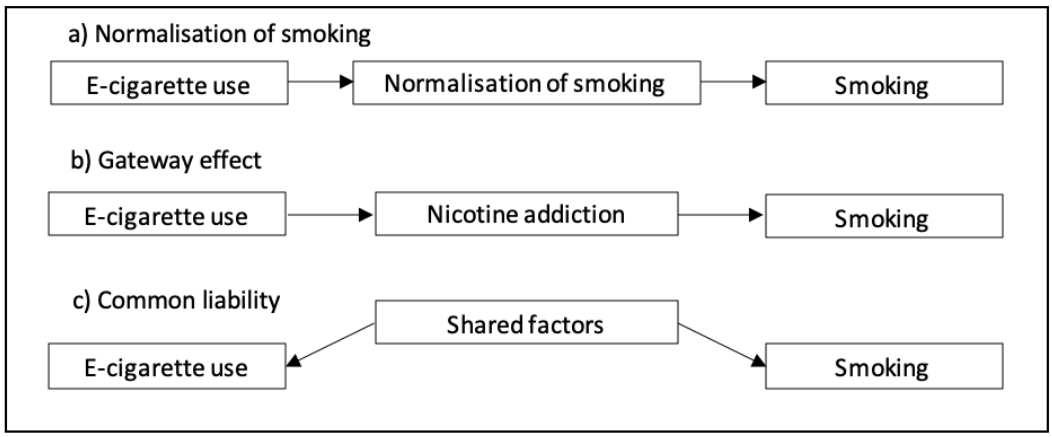


Figure 1.2. Diagrams to show the potential mechanisms (normalisation [a], gateway effect [b] or common liability [c]) of the association between e-cigarette use among never vapers and later smoking.

1.3.3 Normalisation

Concerns have been raised over the potential implications of e-cigarette use on an individual and societal scale (Sæbø & Scheffels, 2017). One such concern is that the presence of e-cigarettes in public areas, particularly where it is illegal to smoke but not to vape, may encourage smoking by normalising the behaviour (Fairchild, Bayer, & Colgrove, 2013; Stanwick, 2015). Arguably, seeing individuals vape in public may normalise nicotine addiction and frame smoking as a socially acceptable behaviour because the behaviours are similar in appearance (i.e., a hand to mouth action followed by a visible exhalation of smoke/vapour). On an individual level, if a non-smoker sees vaping or nicotine addiction as normal and acceptable, they may be more likely to start vaping; they may also see the act of smoking as more normal due to the similarities with vaping and thus be more likely to experiment with smoking too (Sæbø & Scheffels, 2017). However, there is little evidence to support the normalisation of smoking among young people in Great Britain; smoking rates continued to decline after the introduction of e-cigarettes despite a marginal slowing in the rate of decline for regular smoking between 2010 and 2015, when e-cigarettes were widely available but unregulated (Hallingberg et al., 2020). Furthermore, over the same time period there was an acceleration in decline of acceptability with regards to smoking (i.e., young people were increasingly less likely to believe it was OK to smoke). Additionally, qualitative evidence suggests there is limited support for e-cigarette use renormalising smoking among young people in Great Britain (Brown et al., 2020). Given the current evidence suggests that it is unlikely that normalisation of smoking can explain the relationship between e-cigarette use and later smoking, I focus on alternative theories in this thesis.

1.3.4 The Gateway Hypothesis

Another theory which has been proposed as a potential explanation for the strong association between vaping among never smokers and subsequent smoking is that vaping acts as a gateway to smoking (Bell & Keane, 2014; Chapman et al., 2019; Etter, 2018). The gateway hypothesis (as this theory is often called) was originally proposed to

explain patterns of recreational drug use (Vanyukov et al., 2012). The original hypothesis referred to the use of a 'soft' drug (e.g., cannabis) leading to the use of a 'hard' drug (e.g., heroin), with three main principles; 1) the 'soft' drug exposure occurs prior to the use of the 'hard' drug, 2) there is an increased risk of subsequent 'hard' drug use among 'soft' drug users compared to non-users, and 3) there is a dose-response relationship whereby the more an individual uses the 'soft drug', the more likely they are to use the 'hard' drug (Vanyukov et al., 2012). Many have applied this hypothesis to the relationship between e-cigarette use and later smoking whereby e-cigarette use is considered to be the 'soft' drug and cigarettes are considered to be the 'hard' drug. Although both products (can) deliver nicotine, cigarettes are considered 'hard' due to the other constituents of tobacco smoke which are not present in e-cigarettes and due to the known harms of smoking.

Proponents of the gateway effect often claim that the effect is attributable to nicotine addiction (Bell & Keane, 2014). Historically (i.e., among older generation devices), e-cigarettes have not delivered nicotine as effectively as cigarettes (Farsalinos et al., 2014); therefore, some e-cigarettes may not be adequate to satisfy users who become more heavily addicted to nicotine. Consequently, addicted users may transition to smoking in order to satisfy their needs. However, newer generations of e-cigarettes can deliver similar levels of nicotine (Bowen & Xing, 2015), so this hypothesis may not be sufficient to explain why users of newer devices may transition to smoking. Alternatively, these newer generation devices may increase the likelihood of developing nicotine dependence (as they deliver such high doses of nicotine), which may increase the likelihood of users smoking if users need to satisfy their nicotine addiction by any means but cigarettes are more available than e-cigarettes.

Despite a lack of consideration of nicotine contents in the studies included in Soneji and colleagues' (2017) meta-analysis, Soneji and colleagues concluded that there were several aspects of the association between e-cigarette use and later smoking that suggested a causal effect (e.g., a gateway effect). However, Soneji and colleagues (2017) did not stratify by risk of study bias, which was high for many of the included studies (Glasser et al., 2018). This is important because the preconceptions of study authors may also consciously or unconsciously influence how studies are designed and conducted, and this may be reflected in a study's conclusions. For example, two studies in the Soneji and colleagues (2017) meta-analysis drew diverging conclusions, despite

the pooled odds ratios not differing substantially from each other; Leventhal and colleagues (2015) concluded there was insufficient evidence to support the gateway hypothesis, whereas Miech and colleagues (2017) concluded that there was a one-way bridge from e-cigarette use to smoking.

1.3.5 The Common Liability Theory

An alternative theory which is sometimes used to explain the association seen between e-cigarette use (among never smokers) and subsequent smoking is the common or shared liability theory (Etter, 2018). Proponents of the common liability theory suggest that people who use multiple drugs (or in the case of e-cigarettes and cigarettes, different delivery methods of the same drug – nicotine) share the same predisposing factors (Chapman et al., 2019). When considering a causal relationship between e-cigarette use and later smoking (e.g., the gateway hypothesis), these shared factors/confounders (i.e., common liabilities) should be taken into account. Often the studies included in the Soneji and colleagues (2017) meta-analysis, failed to sufficiently adjust for potential confounders (Glasser et al., 2018) meaning the association found may be due to shared factors which influence the likelihood of engaging in both behaviours. I explore this further in Chapter 2.

1.4 Demographics of E-cigarette Users

Given that the relationship between e-cigarette use and smoking could be due to a shared liability, it is vital to observe the demographics of e-cigarette users. Previous evidence indeed suggests that e-cigarette users differ in terms of sex, income, socioeconomic position, smoking history, and age compared to non-users.

1.4.1 Sex

Using nationally representative data from the Tobacco Use Supplement of the Current Population Survey for the United States (US), Levy and colleagues (2017) found a higher proportion of men in the US have ever used e-cigarettes compared with women (8.6% and 7.0% respectively). Men are also more likely to currently (in the past 30 days) and regularly (20 or more times in the past 30 days) use e-cigarettes than women in the US (current use: 2.3% and 1.9% respectively; regular use: 1.1% and 0.8% respectively). Similarly, data from the adult Special Eurobarometer for Tobacco survey suggest that

males in the European Union are also more likely to ever use e-cigarettes than females (Filippidis, Laverty, Gerovasili, & Vardavas, 2017), and in a cross-sectional study in Malaysia, sex differences are observable to a greater degree, where 18.1% of male respondents were e-cigarette users compared to 0.4% of female respondents (Perialathan et al., 2018).

In a systematic review observing differences in e-cigarette use between sociodemographic groups, there were no clear differences between males and females in terms of ever use in two high-quality studies – one observing children in Wales and one observing adults in the US (Hartwell, Thomas, Egan, Gilmore, & Petticrew, 2017). However, four out of five medium-quality studies (conducted in Poland and the US) showed that men were more likely to have ever used e-cigarettes than women and three out of four studies (conducted in the US, Poland and South Korea) showed men were more likely to currently use e-cigarettes (Hartwell et al., 2017). Combined, these results suggest that men may be more likely to vape than women.

1.4.2 Income and Socioeconomic Position

The association between e-cigarette use and income is less clear. For example, US adults who quit smoking more than 3 years ago or have never smoked are more likely to have a *lower* income if they regularly vape, whereas those who have smoked within the past 3 years are more likely to regularly vape if they have a *higher* income (Levy et al., 2017). Among Europeans in the adult Special Eurobarometer for Tobacco survey, there is weak evidence to suggest that ever e-cigarette users are more likely to experience financial difficulties; however, there is no clear evidence to suggest that regular users are more likely to experience financial difficulties (Filippidis et al., 2017). Data from the International Tobacco Control (ITC) Four-Country Survey of adults in the US, Canada, Australia, and the UK showed that higher income was associated with ever vaping (particularly in the US and UK), but not current vaping (Adkison et al., 2013).

Income is often used as a proxy for socioeconomic position, alongside education and occupation (Galobardes, Shaw, Lawlor, & Lynch, 2006a, 2006b). A systematic review observing differences in e-cigarette use by sociodemographic groups found that levels of association between socioeconomic factors and e-cigarette use vary between and within countries (Hartwell et al., 2017). There was limited evidence for an association between

occupation and e-cigarette use, but ever e-cigarette use was frequently associated with higher educational levels (Hartwell et al., 2017). Data from the US Legacy Longitudinal Smoker Cohort Study suggest that those in higher socioeconomic groups (i.e., higher educational level) in the US are more likely to vape, but there was no clear evidence of this association in an online survey (Pearson, Richardson, Niaura, Vallone, & Abrams, 2012).

Socioeconomic disadvantage (i.e., lower socioeconomic position based on educational level, occupational status and net income), is associated with increased likelihood of current vaping among never smoking youth and adult ex-smokers but not adult never or current smokers in the UK (Green, Gray, Sweeting, & Benzeval, 2020). Additionally, evidence from a national cross-sectional UK survey of current adult smokers and adult smokers that recently quit (within the past year) suggests that higher socioeconomic position (assessed using the occupation-based National Statistics Socio-Economic Classification) is associated with higher likelihood of current e-cigarette use among current smokers only (Brown et al., 2014). Where smoking history is taken into account in studies exploring the association between socioeconomic position and vaping, the association appears to be mediated by smoking history (Brown et al., 2014; Green et al., 2020; Levy et al., 2017).

1.4.3 Smoking History

In a cross-sectional study, Levy and colleagues (2017) found that US smokers and ex-smokers who had quit in the last 3 years were more likely to have ever vaped than non-smokers or ex-smokers of more than 3 years. Current and regular vapers were also less likely to be current smokers compared with ever vapers (Levy et al., 2017). Using ITC study data, Adkinson and colleagues (2013) found that non-daily smokers in the US, UK, Australia and Canada were more likely to have tried e-cigarettes, and long-term ex-smokers were less likely to have tried e-cigarettes than daily smokers (≤ 20 cigarettes per day). Heavy smokers (who smoked 20 or more cigarettes per day) and non-daily smokers were more likely to currently vape than respondents who were daily smokers (≤ 20 cigarettes per day). Similarly, in the UK, Brown and colleagues (2014) found that smoking a higher number of cigarettes per day and having attempted to quit smoking in the past year are associated with higher likelihood of current vaping among adults.

As previously discussed, young people in the UK who are ever smokers are also more likely to subsequently vape than never smokers (East et al., 2018). In a longitudinal, 3-wave study of Canadian youth, infrequent and frequent smokers at wave 1 were more likely to be dual users (using both cigarettes and e-cigarettes) at waves 2 and 3 compared to non-smokers (Aleyan, Hitchman, Ferro, & Leatherdale, 2020). Yet Bold and colleagues (2018) used autoregressive cross-lagged models to explore the relationship between current smoking and current e-cigarette use and found no clear evidence that smoking is associated with later e-cigarette use among high school students in Connecticut (US). Importantly, this model did not account for regular use of either product (prior to wave 1 or between waves), which could be problematic if there are age specific factors which influence use of each product differently (i.e., if smoking initiation is more likely to occur at a younger age than vaping and consequently was not captured in this data).

1.4.4 Age

A systematic review observing differences in e-cigarette use by sociodemographic groups found that all high-quality studies showed greater ever and current use of e-cigarettes among older adolescents (17-18 years) and/or young adults (18-24 years) compared with younger children (12-13 years) or older adults (25-39 years) (Hartwell et al., 2017). Similarly, ITC data of adults in the US, Canada, Australia, and the UK also suggested that 18-24 year olds were most likely to try e-cigarettes but there was no clear association between age and current use of e-cigarettes (Adkison et al., 2013). In Great Britain in 2019, current e-cigarette use was most prevalent among 35-44 year olds (9.5%) followed by 45-54 year olds (9.3%), and then 25-34 year olds (7.8%), but only 4.3% of young adults aged 18-24, and 5.6% of those over 55 years currently used e-cigarettes (Action on Smoking and Health, 2019a). Among young people in Great Britain (11-18 years), current vaping has increased from 2.4% in 2015 to 4.9% in 2019 (Action on Smoking and Health, 2019b).

1.5 Reasons for Vaping

There are clear differences in e-cigarette usage patterns across age groups which could indicate different factors influencing likelihood of use. Reasons for vaping, for example,

may differ between adolescents, older adults, and young adults, and these reasons may impact smoking and vaping behaviour.

1.5.1 Adolescents

In a study of South Korean adolescents (13-18 years), Lee, Lee, and Cho (2017) found that the most common reason for vaping among infrequent monthly users was out of curiosity (0-2 days a month = 28.8%), but the most common reasons for more frequent monthly vaping were to quit smoking (3-9 days a month = 18.7%; ≥10 days a month = 21.0%) and to vape indoors (3-9 days a month = 17%; ≥ 10 days a month = 19.5%). Among US middle and high school students, vaping 'out of curiosity' was the most common reason for use (57.1%), followed by 'good flavours' (41.8%) and 'friends use' (32.6%), but using e-cigarettes to quit smoking was the strongest predictor of continued e-cigarette use as these users were 13 times more likely to continue vaping than those who did not use e-cigarettes to quit smoking (Bold, Kong, Cavallo, Camenga, & Krishnan-Sarin, 2016).

1.5.2 Adults

In contrast to the majority of research on adolescents, the primary reasons for vaping among adults (18+ years) in Great Britain are related to smoking cessation (Action on Smoking and Health, 2019a); 21% of respondents to a national survey stated they use e-cigarettes to cut down their smoking, 14% stated they vape to quit smoking, and 12% stated they vape to prevent relapse. Among American Indian adult (18+ years) dual users in Oklahoma, a much larger proportion (79%) vaped to quit smoking (Rhoades et al., 2019).

In the nationally representative Population Assessment of Tobacco and Health survey in the US, use of e-cigarettes as an alternative to cigarettes (a factor created using multiple reasons such as 'using them help people to quit smoking') was found to be associated with current vaping and former regular vaping (Nicksic, Snell, & Barnes, 2019). Similarly, Yong and colleagues (2019) reported results from the ITC study which showed that smokers were more likely to currently vape if they vaped to cut down their smoking (86%), thought it was less harmful to others (78%) or vaped to help them quit smoking (77%). Current smokers who vaped daily were also found to be more likely than non-daily vapers to vape to help with quitting smoking, to cut down their smoking, out

of enjoyment, due to affordability or because they thought it was more acceptable than smoking (Yong et al., 2019). Ex-smokers were more likely to currently use e-cigarettes out of enjoyment (91%), because they thought they were less harmful to others (90%), due to affordability (90%) or to help to stay quit (88%), and were more likely to vape out of enjoyment or due to affordability if they vaped daily rather than weekly (Yong et al., 2019).

1.5.3 Young Adults

Yong and colleagues (2019) also reported some differences between reasons for use and associated behaviour among younger adults (18-24 years); compared with older adults, younger people appeared to be more motivated to regularly vape for reasons other than quitting smoking. Similar to the findings among adolescents, evidence from the US suggests that young adults vape primarily out of curiosity (Kong, Morean, Cavallo, Camenga, & Krishnan-Sarin, 2015), enjoyment (Saddleson et al., 2016) or because their friends or family vape (Tsai et al., 2018).

Although some of the studies described here have explored associations between reasons for vaping and continuation/discontinuation of vaping and smoking, there is limited evidence from young adults in the UK specifically. Understanding why young adults vape, and how this influences their later vaping and smoking behaviour, is necessary to determine whether cautious policy approaches (e.g., bans) are appropriate and could also help to inform public health strategies. I explore this in Chapter 3.

1.6 Genetics and E-cigarette Use

Aside from the subjective reasons given for e-cigarette use, there may be other underlying factors influencing peoples' use of e-cigarettes, such as their genetics.

1.6.1 Genetics

Some people are genetically predisposed to develop certain diseases or engage in certain behaviours (i.e., phenotypes). For example, people with specific mutations in the *BRCA1* and *BRCA2* gene are more likely to develop breast cancer than those without such mutations (Kuchenbaecker et al., 2017). However, most traits are polygenic, meaning they are influenced by multiple genetic variants from a range of genes (Shi,

Kichaev, & Pasaniuc, 2016). For example, 378 genetic variants have been found to be associated with liability for smoking initiation (Liu et al., 2019).

Nucleotides are chemical bases which lie on the DNA strand (Alberts et al., 2002). There are four types of base, Thymine (T), Adenine (A), Cytosine (C), and Guanine (G). Due to their chemical structure, A always pairs with T, and C always pairs with G. In humans, the order of these bases on the DNA strand is 99.9% identical, but 0.1% of these bases can vary. Some of this variation occurs due to substitutions of one base for another (e.g., an A is substituted for a T) at a single base pair. Common genetic variants of this type are called single nucleotide polymorphisms (SNPs). These SNPs can result in differences between humans for a variety of traits, including eye colour, height, disease risk, smoking behaviours, and so on.

1.6.2 Genome-Wide Association Studies

To discover whether individuals are genetically predisposed to display a phenotype, researchers first need to know which SNPs are associated with that phenotype. Genome-wide association studies are designed to discover which genetic variants influence a specific trait and therefore can be used to identify these SNPs. A genome-wide association study can explore the association between over 8 million genetic variants across the genome and a specific phenotype such as eye colour or smoking initiation (Bush & Moore, 2012). The results of a genome-wide association study are often displayed in a Manhattan plot with SNP position along the x-axis, significance level along the y-axis, and a significance line indicating that all values lying above this threshold are considered to be clearly associated with the phenotype, although some influential SNPs may be missed at this highly stringent threshold (Reed et al., 2015). This threshold is usually $p < 5\times10^{-8}$ (known as the genome wide significant threshold) to account for multiple testing of the likely number of functional units in the human genome (Hoggart, Clark, De Iorio, Whittaker, & Balding, 2008; Panagiotou, Ioannidis, & the Genome-Wide Significance Project, 2011).

1.6.3 Polygenic Risk Scores

Individual SNPs identified in genome-wide association studies as associated with traits of interest can be combined into polygenic risk scores. Polygenic risk scores are weighted sums or averages of the number of risk alleles (identified in a genome-wide association

study and weighted by the effect size of each variant on the trait of interest from that genome-wide association study) each individual possesses (Lewis & Vassos, 2020). Polygenic risk scores can be calculated using software such as PRSice (Choi & O'Reilly, 2019; Euesden, Lewis, & O'Reilly, 2015) and PLINK (Chang et al., 2015; Purcell et al., 2007). Higher polygenic risk scores are indicative of greater genetic predisposition to the given phenotype. Once calculated, polygenic risk scores can be used to predict individuals' predisposition to a given phenotype and observe the association between predisposition to a phenotype and outcomes of interest (Lewis & Vassos, 2020).

1.6.4 Genome Wide Association Studies of E-cigarette Use and Smoking Initiation

At present, there are no large-scale genome-wide association studies which could be used to create an e-cigarette use polygenic risk score. Ideally, I would conduct a genome-wide association study to identify SNPs which are associated with vaping, but I do not have access to data with a large enough sample of vapers with available genetic data. Therefore, I am currently unable to observe the association between genetic predisposition to vaping and outcomes such as smoking. However, it is biologically plausible that there could be a genetic overlap between vaping and smoking initiation as both cigarettes and e-cigarettes (can) contain nicotine; a predisposition to smoking initiation could be in part due to an individuals' response to nicotine or their nicotine metabolism.

Allegrini and colleagues (2019) have investigated this overlap; using independently associated SNPs from the Tobacco and Genetic Consortium (TAG) genome-wide association study to create polygenic risk scores for smoking initiation and smoking heaviness, they found that a predisposition to be a heavier smoker was associated with an increased likelihood to ever use e-cigarettes. They concluded that this association may be due to a genetic predisposition to use nicotine or that it could reflect a more general personality trait, such as impulsivity or risk-taking.

Since Allegrini and colleagues (2019) published these findings, a larger genome-wide association study of smoking initiation and smoking heaviness has been published using data from the Genome-Wide Association Study and Sequencing Consortium of Alcohol and Nicotine (GSCAN) which contains data from over 1.2 million individuals (Liu et al.,

2019). The TAG genome-wide association study identified 8 SNPs associated with smoking initiation and 3 SNPs associated with smoking heaviness, whereas the GSCAN genome-wide association study identified 378 and 55 respectively. Thus, using the SNPs identified in the GSCAN genome-wide association study (rather than the SNPs identified in the TAG genome-wide association study) should allow for greater statistical power to detect associations as the polygenic risk scores should explain a greater proportion of variance in the phenotype. In Chapter 4, I have explored the association between smoking initiation polygenic risk scores identified in GSCAN and later e-cigarette use. I focussed on smoking initiation rather than smoking heaviness due to the potential for collider bias and reduced statistical power that would result from stratifying by smoking status in order to explore the associations with smoking heaviness.

1.6.5 Positive and Negative Controls

Allegrini and colleagues (2019) suggested that the association found between smoking polygenic risk scores and e-cigarette use could be due to personality traits such as risk-taking. One way to explore this hypothesis would be to look at the association between smoking initiation polygenic risk and e-cigarette use and compare it with the association between the polygenic risk scores and selected control outcomes.

Control outcomes, which can either be positive or negative, are often used in genetic research and are useful for a variety of reasons. Positive controls can be used to check that polygenic risk scores are indeed associated with the phenotype of interest they should be predicting. The magnitude of association between the polygenic risk score and phenotype of interest (e.g., smoking initiation) can also be compared with the association between the polygenic risk score and the outcome of interest (e.g., e-cigarette use) to explore potential mechanisms underlying the associations. An example of a positive control would be looking at the association between alcohol consumption polygenic risk scores and self-reported alcohol consumption. If there is no clear evidence of association, it would suggest that the polygenic risk score is not adequately predicting predisposition to alcohol consumption in the chosen sample. Negative control outcomes are outcomes for which there is no plausible biological route to the exposure. They can inform the overall evaluation of whether an association is causal via a hypothesised route. For example, smoking is associated with risk of dying by suicide (which is biologically plausible), but equally strongly associated with risk of dying

by homicide (which is not), casting doubt on the causal nature of the former association (Davey Smith, Phillips, & Neaton, 1992). Therefore, associations with negative controls suggest there is another factor explaining the association between polygenic risk score and the outcome (i.e., confounding). The association between the polygenic risk score and the outcome may still be causal, but this is less likely if it has a similar confounding structure to the negative control association.

Comparing associations between smoking initiation polygenic risk scores and e-cigarette use with associations between smoking initiation polygenic risk scores and biologically unrelated risky behaviours, such as risky sexual behaviour, could indicate whether it is possible that the association found by Allegrini and colleagues (2019) could be explained by a general risk-taking disposition. I explore this possibility in Chapter 4.

1.7 Predicting Health Outcomes of E-cigarette Use

Novel methods and supplementary analyses, such as the use of negative control outcomes, are required to aid understanding of the factors influencing e-cigarette use and the possible consequences of use. As e-cigarettes are relatively new, little is known about the long-term health effects of vaping and the effects will not be observable for many years to come. Even when these effects are eventually observable, a high proportion of e-cigarette users have also smoked, meaning any observable effects will be heavily confounded by smoke exposure – a known cause of many poor health outcomes. Despite this issue, some research has attempted to explore the potential health consequences of e-cigarette use.

1.7.1 Potential Health Consequences of E-cigarette Use

E-cigarette users may be exposed to levels of aldehydes (e.g., formaldehyde) which are equal to or surpass levels resulting from smoke exposure (Gillman, Kistler, Stewart, & Paolantonio, 2016; Jensen, Luo, Pankow, Strongin, & Peyton, 2015); levels at which exposure is associated with an increased risk of cancer among current smokers (Godish, 1989). However, these studies are based on chemical analysis of artificial puffing behaviour (i.e., using a smoking machine) and consequently there has been considerable debate about the extent to which e-cigarette users are exposed to formaldehyde. Some researchers claim vapers are only exposed to dangerous levels under conditions which

are unpleasant to users; Farsalinos and colleagues (2015) state that the voltages necessary to produce vapours containing harmful levels of formaldehyde create an unpleasant 'dry puff' sensation. Users of new generation e-cigarettes (which are more powerful than earlier generation e-cigarettes) can adjust the voltage of their e-cigarette, but quickly learn to avoid 'dry puff' sensations caused by excessively high voltages (Farsalinos et al., 2015). However, Salamanca and colleagues (2018) claim that dangerous levels of formaldehyde can be produced at voltages which are non-aversive to users.

Smokers are exposed to numerous carcinogens, but e-cigarette users have been shown to have lower levels of carcinogens in their saliva and urine compared with smokers (Shahab, Dobbie, Hiscock, McNeill, & Bauld, 2017). Although this suggests that smokers who switch to e-cigarettes will be at lower risk of developing cancer, never-smoking e-cigarette users could be exposing themselves to low levels of carcinogens which could have negative effects following repeated exposure (Glantz & Bareham, 2018). However, low levels of carcinogens can be detected at levels unlikely to have any biological effects, so the measurable presence of carcinogens alone is not sufficient to suggest harmful outcomes; it is important to also consider the level of carcinogen exposure (Nohmi, 2018).

Smokers and e-cigarette users are also exposed to ultrafine particles in the smoke/vapour which they inhale. Although inhaling ultrafine particles is associated with poor pulmonary and cardiovascular outcomes (Glantz & Bareham, 2018), many of these associations are evidenced in smokers, and there is no clear evidence that ultrafine particles in e-cigarette vapour will have the same health effects as cigarette smoke. Crucially, the e-cigarette vapour in which the ultrafine particles are carried is chemically and physically different to cigarette smoke, and the ultrafine particles may differ in size, composition and toxicity (Grana, Benowitz, & Glantz, 2014). Therefore, it is unclear whether ultrafine particle exposure via e-cigarette use will impact pulmonary or cardiovascular health.

Chronic obstructive pulmonary disease (COPD) is strongly associated with smoking and evidence suggests that it may also be strongly associated with vaping (Osei et al., 2020; Xie, Ossip, Rahman, & Li, 2019). The NHS claim that 9 out of 10 COPD cases are likely to be a result of smoking, with other possible causes including occupational fumes and

dust, air pollution and genetics, but there is no discussion of e-cigarette use as a potential cause (NHS, 2019a). Osei and colleagues (2020) reported a strong association between current e-cigarette use and COPD, particularly among never smokers (less than 100 cigarettes in their lifetime) who used an e-cigarette daily (OR = 2.64, 95% CI = 1.43 to 4.89). However, the study did not adequately take into account the extent of tobacco exposure; never smokers self-reported smoking less than 100 cigarettes in a lifetime but may have been exposed to tobacco from other sources (e.g., cigars, shisha). Furthermore, due to the cross-sectional nature of the study, direction of causality cannot be inferred, and COPD diagnosis may lead to e-cigarette use. It is unlikely that COPD causes never-smokers to start using an e-cigarette, however, misreporting of smoking behaviour can introduce bias in studies that rely on self-reports (Khouja, Munafò, et al., 2020), and misreporting is a common issue among populations with COPD (Shahab, Jarvis, Britton, & West, 2006). If self-reported never smokers misreport their behaviour (i.e., they have actually smoked more than 100 cigarettes in their lifetime), the wrong direction of causality could be inferred; the observed relationship would indicate that vaping caused the COPD diagnosis, however, smoke exposure may have caused the COPD diagnosis, which in turn caused the individual to switch from smoking to vaping. Xie and colleagues (2019) also showed a positive cross-sectional association between e-cigarette use and self-reported COPD compared with never smokers, but compared with smokers and dual users there was evidence of a *lower* likelihood of self-reported COPD. Longitudinal evidence, albeit with a small sample size and short follow up, does not suggest that e-cigarette use is associated with respiratory symptoms among never smokers who have vaped for at least 3.5 years (Polosa et al., 2017).

Bronchiolitis obliterans (also known as popcorn lung) is a lung disease that is similar to COPD, and associated with inhalation of diacetyl in popcorn factory workers after ~12 months of repeated, short-term, intense exposure (Kanwal et al., 2006; Kreiss et al., 2002). Diacetyl was found in a majority of tested e-cigarettes (39 out of 51 products) in 2016 (Allen et al., 2016). However, there are a lack of cases found in e-cigarette users. Furthermore, the studies which have found associations between diacetyl inhalation and popcorn lung among popcorn factory workers have been heavily criticised for failing to take into account the smoking history of the cohort or the disproportionately high rates of exposure to other possible causes of lung disease, such as farming-related exposures

(Pierce, Abelmann, & Finley, 2014; Pierce, Abelmann, Spicer, Adams, & Finley, 2014). In fact, diacetyl exposure due to smoking far exceeds that of occupational exposure, yet smoking has not been found to be associated with bronchiolitis obliterans (Pierce, Abelmann, Spicer, et al., 2014). Despite the weak evidence for a causal effect of diacetyl on popcorn lung among e-cigarette users, diacetyl was banned from e-cigarettes sold in the UK and Europe with the enforcement of the European Tobacco Products Directive in 2016. However, diacetyl remains in products in other areas of the world such as Canada and the US.

The European Tobacco Products Directive also prevents the inclusion of vitamins such as vitamin E acetate in e-cigarettes. Vitamin E acetate, an ingredient added to tetrahydrocannabinol (THC) products as a thickener, has been identified as the probable cause of e-cigarette or vaping product use associated lung injury (EVALI) (Blount et al., 2019). In the US, 68 individuals have died due to EVALI, and more than 2,500 have been hospitalised (as of February 2020), yet there have been no cases of EVALI in the UK (CDC, 2020). Although the harmful products were primarily acquired from informal, unregulated sources (e.g., friends, acquaintances and unlicenced retailers) rather than licenced vaping shops or online stores (Heinzerling et al., 2020), and no cases were identified in the UK, the EVALI outbreak coincides with an increase in people in the UK believing that e-cigarettes are as harmful as, or more harmful than, smoking (Tattan-Birch, Brown, Shahab, & Jackson, 2020).

Evidence published in 2019 suggesting that e-cigarette use is associated with increased risk of myocardial infarction (i.e., heart attacks) could also have contributed to the increase in harm perception of e-cigarettes (Bhatta & Glantz, 2019). Bhatta and Glantz (2019) explored the association between e-cigarette use and myocardial infarction cross-sectionally and longitudinally but found only a cross-sectional association whereby e-cigarette use increased the risk of myocardial infarction. However, this analysis does not rule out reverse causality as an explanation for the association. To explore whether the association was due to reverse causality, the authors explored whether myocardial infarction predicted daily e-cigarette use at follow up – it did not. However, the study was heavily criticised for not adequately accounting for the order of the exposure and outcome by excluding those who had experienced a myocardial infarction prior to e-cigarette use (which was possible using the available dataset). Following calls for the paper to be retracted, the authors were invited to revise their analysis but declined due

to data access issues. Consequently, the paper was retracted less than a year later. A reanalysis of the study, which adequately accounted for the order of exposure, showed no clear evidence of an association between e-cigarette use and myocardial infarction among never smokers (Rodu & Plurphanswat, 2020). In line with the reanalysis, a review exploring the relationship between electronic cigarette use and cardiovascular health stated there was inconclusive evidence to support a long-term effect on cardiovascular health (MacDonald & Middlekauff, 2019).

1.7.2 Health Effects of Nicotine

Nicotine is an addictive drug which is naturally found in tobacco leaves (Benowitz, Hukkanen, & Jacob, 2009) and is often implicated as a plausible potential cause of poor cardiovascular health related to e-cigarette use (Benowitz & Fraiman, 2017; Kennedy, van Schalkwyk, McKee, & Pisinger, 2019). When nicotine is inhaled via tobacco smoke, it travels from the users lungs to the brain where it binds to nicotinic cholinergic receptors, leading to the release of neurotransmitters and resulting in psychoactive effects (Benowitz, 2010). Exposure to nicotine induces acute effects, including increased heart rate, blood pressure and sympathetic nerve activity to muscle circulation (Najem et al., 2006). Evidence also suggest there may be longer term effects of nicotine exposure; alongside the acute effects on heart rate, nicotine exposure may increase resting heart rate (Linneberg et al., 2015). While multiple reviews have explored the cardiovascular effects of e-cigarette use on cardiovascular health, there is limited evidence to suggest that nicotine is the mechanism by which e-cigarette use may impact cardiovascular health (Benowitz & Fraiman, 2017; Kennedy et al., 2019; Qasim, Karim, Rivera, Khasawneh, & Alshbool, 2017).

At present, there is limited evidence of long-term effects of nicotine use in humans without exposure to the other constituents of tobacco smoke, because other forms of nicotine that were available prior to the introduction of e-cigarettes (e.g., nicotine replacement therapy) were rarely used for long periods of time. Only 6% of those who used nicotine replacement therapy during a quit attempt were still using nicotine replacement therapy one year later (Shahab, Dobbie, et al., 2017). Hajek and colleagues (2019) similarly found low rates of nicotine replacement therapy use at 1 year post-quit date among those assigned to use nicotine replacement therapy in a randomised controlled trial, but found that 80% of those assigned to use e-cigarettes were still using

an e-cigarette. In a UK adult sample of never smokers (i.e., those without exposure to other constituents of tobacco smoke), long-term use of nicotine is even more rare; less than 0.1% of respondents had used nicotine replacement therapy for 12 months or more and only 0.1% had used e-cigarettes for 12 months or more (Jackson et al., 2019). Therefore, it is difficult to observe the long-term health effects of nicotine use without tobacco smoke exposure in humans.

1.7.3 Univariable Mendelian Randomisation

It is difficult to conclude that e-cigarettes cause any health effects with the current available evidence for multiple reasons. In addition to the lack of observational evidence and high potential for confounding, conducting a randomised controlled trial (the gold-standard method for causal inference) would be unethical; exposing nicotine-naïve individuals to e-cigarettes over a long period of time in order to assess whether they develop poor health outcomes would mean knowingly exposing individuals to toxicants and chemicals which they may not otherwise have been exposed to. Thus, methods which have not previously been applied to this research area, such as Mendelian randomisation, are needed to explore potential health outcomes.

Mendelian randomisation is a method often used to infer causality, particularly where a randomised controlled trial would be unethical or impossible to conduct (Davey Smith & Ebrahim, 2003). It is based on the laws of Mendelian genetics: segregation and independent assortment. The law of segregation states that each individual possesses pairs of alleles for each trait and each parent passes only one of those alleles from each pair onto their offspring (Griffiths, Miller, Suzuki, Lewontin, & William, 2000). Mendel's law of independent assortment states alleles are independently assorted at meiosis so that each allele passed down from a parent to their offspring is independent of the other alleles which are passed down (Griffiths et al., 2000). Mendelian randomisation assumes that these laws are held at a population level (i.e., a random assortment of genes are transferred from parents to their offspring) (Nitsch et al., 2006). Although this is not strictly true for the inheritance of genes in homologous chromosomes where linkage disequilibrium (the non-random assortment of genes in close proximity) may occur, it is generally true with respect to the inheritance of genes in non-homologous chromosomes (Lawlor, Harbord, Sterne, Timpson, & Davey Smith, 2008).

In Mendelian randomisation, an individual's genotype or genetic liability – which should not be associated with the confounding factors that often distort observational evidence – is used as an instrumental variable to assign them to a phenotype group (e.g., smoker or non-smoker). Consequently, the instrumental variable used as the exposure (i.e., an individual's genotype or genetic liability) always precedes the outcome as it is determined at conception. This genetic variation mimics the randomisation process in a randomised controlled trial, reducing issues of both confounding and reverse causality (Nitsch et al., 2006; Swanson, Tiemeier, Ikram, & Hernán, 2017). The relationship between the genotype and the outcome is then assessed to estimate the total causal effect of one exposure on one outcome using one or two datasets. Mendelian randomisation often uses one dataset with individual-level data (i.e., a dataset which contains genetic variants, exposure and outcome data for all participants), but summary-level data can also be used on random subsets of the same sample (Lawlor, 2016). Summary-level data contain only the summarised genetic associations with the exposure and outcome. Mendelian randomisation can also be conducted using summary-level data from two separate datasets (Lawlor, 2016).

As shown in Figure 1.3, the basic Mendelian randomisation method is reliant on the instrumental variables (G) included in the model being valid. For an instrumental variable to be valid, there are three key assumptions (Martens, Pestman, de Boer, Belitser, & Klungel, 2006):

Assumption 1 (relevance) – The genetic variant (G) is associated with the exposure (X).

Assumption 2 (exchangeability) – The genetic variant (G) is independent of any measured and unmeasured confounders of the exposure-outcome relationship (U).

Assumption 3 (exclusion restriction) – The genetic variant (G) is only associated with the outcome (Y) through the exposure (X).

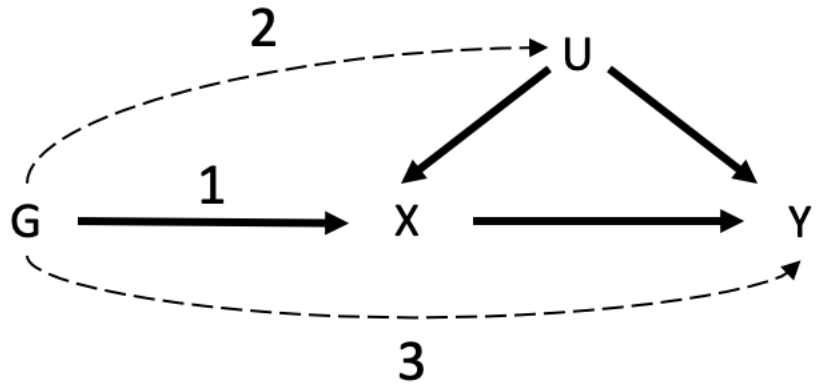


Figure 1.3. A directed acyclic graph depicting the relationship between genetic variant(s) (G), a phenotypic exposure (X), outcome (Y) and confounders (U) in a Mendelian randomisation analysis.

Note: Numbers indicate the three assumptions which valid instrumental variables satisfy. The dashed lines indicate the absence of an association.

Assumptions 2 and 3 can pose an issue in Mendelian randomisation analysis because genetic variants can be pleiotropic – one genetic variant may affect multiple phenotypes. Pleiotropy can be vertical (affecting a phenotype on the causal pathway from exposure to outcome) or horizontal (affecting a different phenotype which affects the outcome through a pathway other than through the exposure or directly affecting the outcome). Although vertical pleiotropy is not problematic for Mendelian randomisation, horizontal pleiotropy can violate assumptions 2 and 3 (Davies, Holmes, & Davey Smith, 2018). Confounding by population structure (i.e., common sub-population differences in allele frequencies) can also violate assumption 2. Various Mendelian randomisation methods have been developed which make different assumptions about pleiotropy and are often used in parallel and compared; consistent estimates across multiple Mendelian randomisation methods (which make different assumptions) provides stronger evidence for a causal effect (Bowden, Davey Smith, Haycock, & Burgess, 2016). For example, the inverse variance weighted method is well-powered but cannot handle any invalid instruments (i.e., violation of assumptions 1, 2 or 3), whereas the weighted median approach can handle violation of assumptions 2 and 3 provided at least 50% of the SNPs are valid instruments. Similarly, the MR-Egger approach allows for

100% of the SNPs used to be invalid due to violation of assumption 3, but an additional and untestable assumption has to be met: the Instrument Strength Independent of Direct Effect (InSIDE) assumption. The InSIDE assumption is a relaxed version of the exclusion restriction assumption (assumption 3) whereby the SNP-exposure effect should not be correlated with any horizontal pleiotropic effect (Bowden, Davey Smith, & Burgess, 2015).

1.7.4 Mendelian Randomisation Studies of Smoking and Health

Although there are no Mendelian randomisation studies exploring the causal effects of vaping, previous Mendelian randomisation studies have found that increased smoking heaviness causes poor lung function, increased risk of COPD (Millard, Munafò, Tilling, Wootton, & Smith, 2018), and increased resting heart rate (Åsvold et al., 2014) as well as increased risk of all-cause mortality, circulatory diseases, and respiratory diseases (Vie et al., 2019). Heavier smoking also appears to cause lower body mass index (BMI) (Åsvold et al., 2014; Taylor et al., 2019). In fact, bidirectional effects of smoking and BMI have been observed; heavier smoking decreases BMI and higher BMI causally increases smoking heaviness (Taylor et al., 2019). However, increased BMI does not appear to causally influence cotinine levels (a highly specific direct metabolite of nicotine), which indicates a more complex relationship between nicotine and BMI.

Although these univariable Mendelian randomisation studies provide evidence for a causal effect of smoking, they cannot determine which constituent of tobacco smoke causes these effects. On the one hand, if nicotine is the constituent driving these effects, I could infer that long-term nicotine use via e-cigarettes would also lead to these poor health outcomes. On the other hand, if nicotine is not driving these effects, I could infer that long-term use of nicotine via e-cigarettes may not lead to these poor health outcomes. Using a multivariable Mendelian randomisation approach, the direct causal effects of nicotine exposure can potentially be determined.

1.7.5 Multivariable Mendelian Randomisation

Multivariable Mendelian randomisation is an extension of the inverse variance weighted Mendelian randomisation method. Rather than calculating the total effect of one exposure on one outcome (which includes indirect effects that mediate the effect of the exposure on the outcome), multivariable Mendelian randomisation is used to explore

the direct causal effect of two or more exposures on an outcome (Lawlor et al., 2019; Sanderson, Davey Smith, Windmeijer, & Bowden, 2019). When two exposures are related, multivariable Mendelian randomisation can estimate the effect of one exposure on an outcome while accounting for the effect of the other exposure on the outcome. Figure 1.4 demonstrates the inclusion of genetic variants (G_1 and G_2) as instrumental variables for two exposures (X_1 and X_2), including those variants which predict both exposures (G_{12}), in a multivariable Mendelian randomisation analysis.

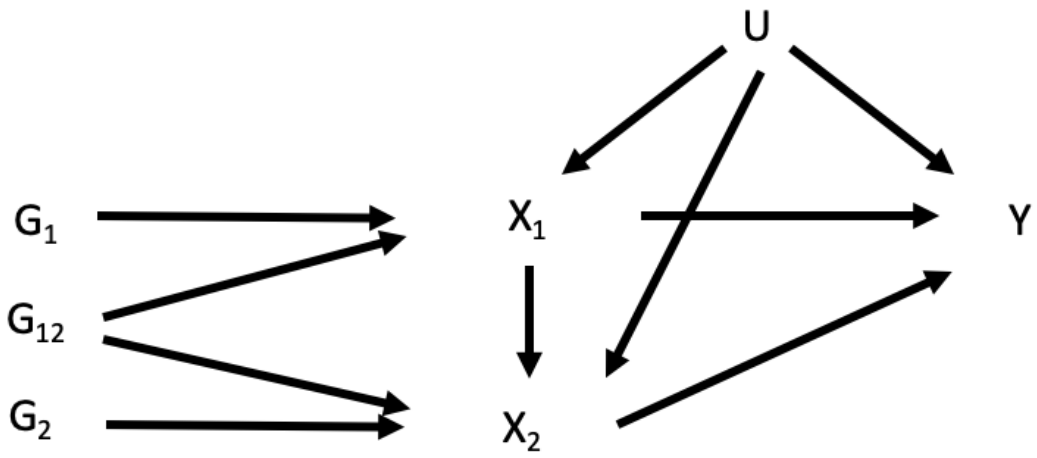


Figure 1.4. A directed acyclic graph to show the relationship between genetic instruments (G_1, G_{12}, and G_2), exposures (X_1 and X_2), confounding factors (U) and outcome (Y) in a multivariable Mendelian randomisation analysis.

Adapted from Sanderson et al. (2019).

1.7.6 Total Versus Direct Effects of Nicotine and Smoking Heaviness

Multivariable Mendelian randomisation could prove particularly useful when investigating the potential health impact of long-term e-cigarette use. As previously discussed, one constituent of particular concern in e-cigarette vapour is nicotine, which is also found in cigarette smoke, but e-cigarettes contain far fewer toxicants which can cause harm. Genetic variants associated with cotinine (a direct metabolite of nicotine) and smoking heaviness have previously been identified in genome-wide association studies (Liu et al., 2019; Ware et al., 2016) which could be utilised in Mendelian randomisation analyses.

Evidence of the health effects of long-term nicotine use is often based on tobacco smoke exposure, but multivariable Mendelian randomisation methods can be used to explore the direct effect of nicotine while taking into account the effect of smoke exposure and vice versa. The total effects of smoking heaviness on health outcomes includes the effect of nicotine on health outcomes, but by exploring the direct effect of smoking heaviness while controlling for the direct effect of nicotine (and confounding factors), it may be possible, in theory, to observe the effect of the remaining constituents of tobacco smoke aside from nicotine. Therefore, among smokers it may be possible to identify the health effects directly caused by nicotine and compare these with the health effects directly caused by the other constituents of tobacco smoke exposure (aside from nicotine) which are less likely to be found in e-cigarette vapour. However, large sample sizes are needed to explore genetic effects and older samples are required to explore long-term health effects. Using the summary results of a genome-wide association study of cotinine levels (as a proxy for nicotine exposure) and another genome-wide association study of smoking heaviness (Liu et al., 2019; Ware et al., 2016), I explore this in UK Biobank (a population-based health research resource consisting of approximately 500,000 people, aged between 38 years and 73 years) in Chapter 5.

1.8 Exploring the Acute Effects of Nicotine

Potential health effects of nicotine use via e-cigarettes can be further explored using experimental methods. Before e-cigarettes became widely available, the short-term effects of nicotine (without the harmful constituents of tobacco smoke) were explored

using nicotine replacement therapy. Exploring short-term effects can provide an insight into the mechanisms by which nicotine could impact long-term health.

1.8.1 Nicotine Use without Exposure to Other Constituents of Tobacco Smoke

Experimental studies have been used to explore the short-term effects of nicotine use. For example, exposure to nicotine replacement therapy acutely increases heart rate by 10 to 15 beats per minute and increases blood pressure by 5 to 10 mm Hg (Benowitz & Gourlay, 1997). These effects can lead to reduced blood flow and oxygen to the heart and result in poor cardiovascular outcomes (Benowitz & Gourlay, 1997), therefore understanding these short term outcomes gives an indication of the potential long-term outcomes.

1.8.2 BMI, Appetite and Nicotine

As previously discussed, evidence suggests there is a complex relationship between BMI and nicotine exposure (Taylor et al., 2019). Observationally, there is a clear association between smoking and BMI; the average BMI of an adult smoker in the US was found to be 2 kg/m^2 lower than the average BMI of a non-smoker, and smoking cessation was associated with 3 to 4 kg of weight gain (Audrain-McGovern & Benowitz, 2011; Williamson et al., 1991). However, Mendelian randomisation studies suggest a bidirectional effect, whereby heavier smoking decreases BMI, but higher BMI increases the likelihood of someone smoking and the number of cigarettes they smoke (Taylor et al., 2019). This could be due to those with higher BMI choosing to smoke due to the known (and frequently cited) association between smoking and lower BMI. There is also some evidence to suggest that e-cigarette use is associated with BMI, but the evidence is limited and the direction of this association is unclear; sole use of either e-cigarettes or cigarettes was found to be associated with high body mass index, but dual use was associated with lower body mass index among adolescents (Jacobs, 2018). One commonly proposed mechanism by which smoking (and potentially e-cigarette use) may impact BMI is through changes in appetite; the NHS website, for example, states that smoking suppresses your appetite, and explains that this effect causes smokers to put on weight when they quit (NHS, 2019b). There is considerable evidence from animal models to suggest that nicotine decreases appetite and feeding (Jo, Talmage, & Role,

2002; Mineur et al., 2011), however, there is limited and conflicting evidence from human studies.

Pilhatsch and colleagues (2014) explored the effect of nicotine versus placebo gum administration on subjective ratings of appetite (measured via visual analogue scale) among human non-smokers after a 30 minute absorption period. Following nicotine administration, non-smokers' subjective ratings of appetite declined by 17.5%. In contrast, non-smokers' appetite ratings *increased* by 18% following placebo administration. Similarly, Perkins and colleagues (1991) found that male participants consumed less food following nicotine administration compared with placebo administration via nasal spray. Although there was no overall difference in food consumption between smokers and non-smokers, further investigation revealed that nicotine administration only clearly reduced food consumption among non-smokers, not among daily smokers. Perkins and colleagues (1992) further explored this effect in another study of male and female daily smokers (using similar methods). Among male and female smokers, caloric intake *increased* following nicotine administration compared with placebo administration. These contrasting results suggest that the effect of nicotine on appetite and food consumption is not as straightforward as it is sometimes claimed to be.

1.8.3 Improving Experimental Designs to Explore Nicotine and Eating Behaviour

There are clear limitations in the existing literature exploring nicotine and eating behaviour. The few available studies in the area are likely to be statistically underpowered; the sample sizes for the studies were small, with as few as 20 participants (Perkins et al., 1992; Perkins et al., 1991). Also, the dose and delivery method of nicotine may not have been appropriate; the studies provided little evidence that the dose was sufficient to induce a biological response among smokers (who are tolerant to nicotine), or whether the method of administration (i.e., nasal spray or gum) led to any side effects, such as changes in taste or sensation, which could affect eating behaviour. Additionally, the contrasting results between smokers and non-smokers suggest that confounding factors were not adequately addressed in the experimental design. For example, nicotine-naïve non-smokers are likely to experience adverse effects of nicotine which could impact their eating behaviour; non-smokers have not developed a tolerance to nicotine whereas daily smokers will have developed a tolerance to

nicotine. This could result in an effect of nicotine administration on appetite and food consumption which is not seen among those who are experienced with nicotine. As well as developing a tolerance to nicotine, daily smokers become dependent on nicotine, thus daily smokers may experience withdrawal symptoms when asked to refrain from smoking. In these studies, smokers were asked to abstain from smoking before the study (in order to standardise nicotine exposure) which could have led to withdrawal symptoms which affected their appetite and eating behaviour (e.g., increased appetite) (Jorenby et al., 1996).

With these limitations in mind, improved experimental studies exploring the effect of nicotine on eating behaviour are necessary before concluding that a nicotine-induced reduction in appetite can explain the relationship between nicotine exposure and BMI. For example, exploring the effects of nicotine on eating topography among non-dependent smokers could avoid the confounding effects of withdrawal seen among dependent smokers and aversion seen among non-smokers; non-dependent smokers are experienced with nicotine and therefore should not experience adverse effects when exposed to nicotine and should also not experience any withdrawal symptoms when asked to refrain from smoking. I explore this further in Chapter 6. If the resulting evidence from such studies supports a nicotine-induced reduction in appetite and eating behaviour, this would support evidence that e-cigarettes may prevent post-cessation weight gain (Russo et al., 2016; Russo et al., 2018). As weight gain following smoking cessation often leads to relapses in smoking (Mizes et al., 1998), smoking cessation services need to understand if there are mechanisms by which e-cigarettes could prevent this, particularly as e-cigarettes are available with and without nicotine.

1.9 Thesis Aims, Methods and Original Contribution

1.9.1 Aims

The aims of this thesis are three-fold, namely to:

1. Identify which young adults are more likely to use e-cigarettes by describing the demographics and profiles of young e-cigarette users in a UK sample;
2. Explore why young adults use e-cigarettes, including the subjective reasons given by vapers, and the genetic factors influencing vaping;

3. Investigate the potential consequences of e-cigarette use on smoking behaviour and health outcomes.

Where possible, I focus on young adults (18 to 30 years), a subgroup who are often understudied despite apparent differences in their motivation and behaviour compared with adolescents and older adults.

1.9.2 Methods

This thesis begins with a systematic review and meta-analysis exploring whether e-cigarettes act as a gateway to smoking (Chapter 2). This review highlights the importance of understanding who uses e-cigarettes, why, and what the potential consequences of use are. In Chapter 3, I first explore the demographics and profiles of young adult vapers in the UK. I go on to explore how young adults' reasons for vaping associate with their later smoking and vaping behaviour using logistic regression. In Chapter 4, I further explore why young adults use e-cigarettes by looking at how smoking initiation polygenic risk scores associate with e-cigarette use, risk-taking and impulsivity in a series of logistic regressions. In Chapter 5, I use univariable and multivariable Mendelian randomisation methods to begin to explore the potential health effects of long-term e-cigarette use by separating the effects of long-term nicotine use from the effects of other constituents of cigarette smoke in UK Biobank. In Chapter 6, I employ an experimental design to further explore the potential health effects of nicotine without exposure to the other constituents of tobacco smoke. Specifically, I explore the effects of nicotine on eating behaviour.

1.9.3 Original Contribution

At present very little is known about which young adults in the UK use e-cigarettes and why, and even less is known about the potential implications of vaping. Using traditional methods, such as randomised controlled trials, to address these research questions would be unethical, therefore alternative methods are required. This thesis contributes a necessary update to existing reviews – which are fast outdated in this rapidly evolving field – and provides a rare insight into the motivations for vaping and self-reported smoking and vaping behaviour of young adults in the UK. Additionally, I have employed novel and innovative methods (e.g., negative control analyses and multivariable Mendelian randomisation methods) to explore research questions which are difficult to

assess observationally at present due to confounding and a lack of longitudinal data. Finally, I provide an outline for experimental research to explore the mechanisms underlying the relationship between nicotine and BMI which is needed to confirm the often cited (but not sufficiently evidenced) claim that nicotine suppresses appetite and food consumption.

Chapter 2 Do e-cigarettes act as a gateway to smoking?

This chapter closely resembles sections from the following publication, pre-print and pre-registered study protocol:

Khouja, J. N., Suddell, S. F., Peters, S. E., Taylor, A. E., & Munafò, M. R. (2020). Is e-cigarette use in non-smoking young adults associated with later smoking? A systematic review and meta-analysis. *Tobacco Control,* Published Online First: 10 March 2020. doi: 10.1136/tobaccocontrol-2019-055433.

Khouja, J. N., Suddell, S. F., Peters, S. E., Taylor, A. E., & Munafò, M. R. (2020). Is e-cigarette use in non-smoking young adults associated with later smoking? A systematic review and meta-analysis. *MedRxiv***;** doi:10.1101/19007005 %J**.**

Khouja, J. N., Suddell, S., Peters, S. E., Taylor, A., & Munafò, M. R. (2019). Does e-cigarette use in non-smoking young adults act as a gateway to smoking? A systematic review and meta-analysis. *OSF*; https://doi.org/10.17605/OSF.IO/3GC2Y.

I took the lead in designing, implementing and interpreting the systematic review and meta-analysis with the support of my co-authors. Specifically, MM and AT provided feedback on the design of the study. Co-authors SS and SP assisted me with study selection and data extraction by independently checking 100% of my selection and extraction (50% each). I was responsible for the data analysis and was responsible for writing the first draft of the manuscript and editing the manuscript in response to comments from co-authors and reviewers. All co-authors provided feedback during the manuscript preparation.

2.1 Chapter Overview

In this Chapter, I describe a systematic review and meta-analysis of the studies investigating the association between e-cigarette use among non-smokers and later smoking. As discussed in Chapter 1, this is not the first review or meta-analysis to address this research question (Glasser et al., 2018; Soneji, Barrington-Trimis, Wills, Levanthal, et al., 2017) but considering my prior knowledge of the evidence base and

that this is a fast-moving area of research, I expected to identify a substantial number of studies that had been published since the previous reviews. As well as providing an update to the currently available reviews, I outline the key areas in which future studies could focus in order to better address the research question. Outlining these areas for improvement helped to shape the design of the studies in the subsequent chapters of my thesis.

2.2 Introduction

There are concerns, shared by the public and policy makers alike, that e-cigarettes may act as a gateway to smoking cigarettes among young people (Chapman et al., 2019). If e-cigarettes act as a gateway to smoking, rather than seeing a decline in smoking rates due to smokers using e-cigarettes for smoking cessation, we may see smoking rates remaining stable or even increasing due to a new generation of smokers for whom e-cigarettes have acted as a route into smoking. This hypothesis is sometimes referred to as the 'gateway hypothesis' or 'catalyst model' and has been widely debated among researchers and public health officials (Etter, 2018). Some argue that rather than e-cigarette use acting as a gateway to smoking, an association between e-cigarette use and later smoking may be explained by a common liability between vaping and later smoking, whereby the same genetic or environmental factors that increase the likelihood of someone vaping also increase the likelihood of someone smoking (Etter, 2018). Such factors should be accounted for when exploring the association between e-cigarette use and later smoking. Furthermore, as many people use e-cigarettes to help them stop smoking (Action on Smoking and Health, 2017), it would also be logical to assume the opposite direction of causality with smoking causing people to vape (Eastwood et al., 2015). The lack of consensus on the issue demonstrates the need for the current available evidence to be synthesised.

As discussed in Chapter 1, a recent meta-analysis concluded that e-cigarette use is associated with a nearly 4-fold increase in the likelihood of smoking at follow up (Soneji, Barrington-Trimis, Wills, Levanthal, et al., 2017). Given this is a fast-moving field, and the topic is of great interest to researchers and policy makers, regular updates to meta-analyses in this area are likely to be necessary. Moreover, in the previous meta-analysis (Soneji, Barrington-Trimis, Wills, Levanthal, et al., 2017), moderate heterogeneity was observed between the study results (I^2 = 60). Some potential sources

of heterogeneity could include the age range of the participants in the studies, cultural and regulatory differences between study locations, and risk of bias among the studies. Soneji and colleagues (2017) addressed one of these potential sources of bias by stratifying by average age, finding that there was greater heterogeneity between studies of adolescents (under the age of 18 years) compared with studies of young adults. However, they did not stratify by risk of bias. This is important because the preconceptions of study authors may also influence how studies are designed and conducted, and this may be reflected in a study's conclusions (as described in Chapter 1).

In this systematic review and meta-analysis, I have updated and extended previous reviews in which e-cigarette use and later smoking have been explored. My aim was to investigate whether e-cigarette use, compared with non-use, in young non-smokers is associated with subsequent cigarette use. I combined evidence from studies investigating this relationship where an odds ratio (OR) could be calculated. Additionally, I explored potential sources of heterogeneity and bias by reviewing the details of included studies, as well as subgrouping, and stratifying the data.

2.3 Methods

The protocol for this systematic review and meta-analysis was pre-registered and published online prior to initiating the systematic search. The pre-registration can be found on the Open Science Framework (https://osf.io/3gc2y/). Where appropriate, I followed PRISMA and MOOSE reporting guidelines.

2.3.1 Eligibility Criteria

In my systematic search of the literature, I included randomised controlled trials, longitudinal studies, cross-sectional studies, and case-control studies and I only included studies investigating young people aged up to 30 years old (inclusive). I selected this cut-off to include studies of both youth and young adults. As a requirement for inclusion, all studies had to have a baseline or retrospective measure of e-cigarette use (including but not limited to ever, occasional, heavy, recent, regular or frequent use) prior to initiating smoking as well as a measure of cigarette smoking (including but not

limited to ever, occasional, heavy, recent, regular, frequent or escalated smoking) as an outcome. Studies also had to include a comparison group (i.e., group which the exposed group is compared to), which could include young people who were never users, trial users (i.e., vaping once/a couple of times or less than 100 times) or not recent e-cigarette users or smokers, dependent on the study. Review articles and animal studies were excluded.

2.3.2 Information Sources

My search strategy was a replication and extension of a strategy used by Soneji and colleagues (2017); I added a search term to the original search criteria, and widened the search by including studies for which an OR could be calculated (rather than where an OR was reported). I conducted an electronic search of the PubMed, Embase, Web of Science and Wiley Cochrane Library databases and also searched the conference abstract books for the Society for Research on Nicotine and Tobacco and the Society for Behavioural Medicine conferences. Due to member restricted access, I was unable to search the NIH Tobacco Regulatory Science Conference abstracts as stated in my pre-registered protocol. I compared the list of studies to be included to those included in previous similar reviews to ensure the amended search strategy had not omitted any relevant studies. Studies written in languages other than English were translated using Google translate where translations were not already available. The search strategy was conducted up to 24th November 2018. E-cigarettes are a relatively new product on the consumer market; therefore, no date restrictions were placed on the search strategy.

2.3.3 Search Strategy

Studies were initially selected for screening using the following search terms within the titles, abstracts or keywords: Cigar* OR Tobacco OR Smok* AND Electronic Cigarette* OR E-Cig* OR Electronic Nicotine Delivery System* OR Vape OR Vaping OR Alternative Nicotine Delivery System*. Boolean operators and truncations differed depending on the database. Relevant medical subject headings (MeSH) terms were included when searching the PubMed database. The search dates and terms used for each platform can be found in Appendix 1. The most notable difference between the search strategy of Soneji and colleagues (2017) and my search strategy is that I included the phrase

'Alternative Nicotine Delivery System'. This term was being frequently used at conferences and in research articles, so I thought it relevant to include in the search.

2.3.4 Study Selection and Data Collection Process

Study selection and data extraction took place over three stages. Stage 1 consisted of title and abstract screening; Stage 2 consisted of a full text screening; Stage 3 consisted of data extraction from selected studies. For each paper, my review team (SS and SP) and I extracted administrative details, study details and participant characteristics. Specifically, these included: author names; year of publication; country of the study; study design; study name (if applicable); sex of included participants, percentage of males included in the total sample and in the case and control groups; number of cases, controls and the size of the cohort; year(s) of data collection; age of the total sample, cases and controls; follow up length (if applicable); comparison group; exposure; outcome; covariates; definition of e-cigarette use and smoking; and type of assessment of e-cigarette use and smoking. We also extracted exposure and control details, outcome details, and results and conclusions. Specifically, these included: stratification information; direction of effect; effect estimate reported; number of individuals included in specific analyses; number of individuals exposed and unexposed in the analysis and number of subsequent smokers for each group; effect size, confidence intervals, standard errors and *p*-values for both unadjusted and adjusted analyses; and the conclusion regarding support for the gateway hypothesis. Titles, abstracts and full text articles were double screened and then double extracted by the review team (myself [100%], SS [50%] and SP [50%]). Discrepancies between myself and the second reviewer were resolved by the third reviewer where necessary. Covidence (www.covidence.org), an online systematic review tool which is in partnership with Cochrane, was used to streamline and document this process.

When insufficient information was available to determine eligibility, I contacted study authors directly via email. Contacting study authors proved challenging, with some authors changing institutions – and consequently changing email addresses – and some being slow to reply or failing to reply at all. Where insufficient information was provided or obtained, the text was excluded from the review.

2.3.5 Risk of Bias Assessment

Risk of bias within studies was assessed using the Newcastle-Ottawa Scale (NOS) (Wells et al., 2012). The selection, comparability and outcome domains of the tool were used to assess the risk of bias in all full texts included in the review (Appendix 2). The studies were rated as good, fair or poor quality based on the star system of the tool (maximum of 9 stars). Thresholds were applied to convert the NOS for study quality to Agency for Health Research and Quality standards (McDonagh, Peterson, Raina, Chang, & Shekelle, 2013) (whereby a good quality rating indicates low risk of bias and a poor rating indicates high risk of bias). Good quality ratings were determined by 3 or 4 stars in the selection domain and 1 or 2 stars in the comparability domain and 2 or 3 stars in the outcome/exposure domain. Fair quality was determined by 2 stars in the selection domain and 1 or 2 stars in the comparability domain and 2 or 3 stars in the outcome/exposure domain. Poor quality was determined by 0 or 1 star in the selection domain or 0 stars in the comparability domain or 0 or 1 stars in outcome/exposure domain. Quality/risk was double assessed by the review team. Studies were not excluded based on their risk of bias, but this information was used to explore heterogeneity by subgrouping.

Risk of bias across studies was assessed using the symmetry and 95% confidence region of a funnel plot (Higgins, Green, & Cochrane Collaboration, 2008). Asymmetry and more than 5% of points lying above the 95% confidence region may indicate some bias across studies.

2.3.6 Causality Assessment

Nine Bradford-Hill criteria are often used to assess causality (Hill, 1965). I selected four of these criteria to indicate the strength of evidence of a possible causal relationship between our exposure and outcome: strength of association, specificity, temporality and dose responsivity. I chose these criteria because they are particularly relevant to studies assessing whether e-cigarettes may act as a gateway to smoking (Etter, 2018) and could be measured on a study-by-study basis (unlike consistency). A strong association was defined as having an adjusted OR of two or more. Studies were considered specific if they adjusted for relevant covariates (e.g., peer smoking/vaping and impulsivity) in addition to basic demographics (i.e., sex, age, socioeconomic position). The temporality

criterion was met if studies were assessed longitudinally (i.e., e-cigarette use was measured at time point 1 with a measure of smoking prior to measuring later smoking at time point 2) – retrospective measures did not meet this criterion. Studies which measured and took into account frequency of e-cigarette use, length of time the product was used for, or how much nicotine was in the e-liquid used, were considered to meet the dose responsivity criterion.

2.3.7 Summary Measures

Effect estimates were reported as ORs (and converted where necessary). ORs of the association between e-cigarette use and later cigarette use were combined using a random-effects model. I chose a random-effects model to account for potential variations in the true effect between studies (e.g., the effect may vary by age thus the true effect may be greater in studies with younger participants). All unadjusted ORs were calculated using observed data points which were obtained from the original study or directly from the author if insufficient information was provided in the original study. Calculated effect sizes were double checked by the review team.

As I did not have access to the original data sets, the adjusted ORs were reported as they were in the original study. Where adjusted risk ratios had been reported, they were converted to ORs using a modified version of a formula published in Section 12.5.4.4 of the Cochrane Handbook (Higgins et al., 2008): OR = (-RR + RR × ACR) / (RR × ACR - 1), where OR = odds ratio; RR = risk ratio; ACR = assumed control risk. I calculated ACR on a per study basis as the risk of later smoking among controls. Sufficient information was available to calculate the ACR for each of the included studies.

2.3.8 Synthesis of Results

I identified five subgroups of interest to explore the data by exposure and outcome definition:

1) ever versus never e-cigarette use at baseline and ever versus never smoking at follow up.

2) ever versus never e-cigarette use at baseline and current (past 30-day) versus non-current use of cigarettes at follow up.

3) current versus non-current e-cigarette use at baseline and ever versus never smoking at follow up.

4) current versus non-current e-cigarette use and current versus non-current use of cigarettes at follow up.

5) regular (at least monthly for more than 6 months) versus non-regular e-cigarette use at baseline and regular cigarette use versus non-regular cigarette use at follow up.

In retrospective studies, measures of e-cigarette use prior to smoking were treated as baseline and smoking status at the time of the study was treated as the follow up.

For the main analysis, I focussed on subgroup 1; I compared ever e-cigarette users with never e-cigarette users at baseline and calculated the pooled ORs (from unadjusted and adjusted ORs) for ever (versus never) cigarette use at follow up. Where multiple results were reported in the original study using varying definitions of the exposure or outcome (e.g., where the results for both subgroup 1 and subgroup 2 were reported), the estimate relating to the comparison of ever versus never e-cigarette use at baseline (subgroup 1) was used in the main analysis. However, if ever use of e-cigarettes (subgroup 1) was not reported, the main effect reported in the study (e.g., subgroup 2) was included in my main analyses.

Pooling estimates from studies with different definitions of e-cigarette use (as I have in the main pooled estimate) can increase risk of bias (Haidich, 2010). Where possible (i.e., where more than one study was available), I also analysed the results of each exposure and outcome definition subgroup. I originally aimed to pool the results of subgroup 5 (comparing regular use); however, insufficient data were available to do so.

Heterogeneity of study effect estimates can be indicated by an I^2 statistic (Higgins et al., 2008). Sources of heterogeneity were explored through subgroup analysis. Specifically, I grouped the studies by risk of bias (poor, fair or high quality), age of participants (including versus excluding participants under the age of 18 years), and location of study (country). All analyses were conducted using Stata SE version 15.1 and Review Manager version 5.

2.4 Results

2.4.1 Study Selection

Figure 2.1 shows the PRISMA study selection flow chart. A total of 15,519 studies were selected for title and abstract screening, 9,199 remained after exclusion of duplicates. After title and abstract screening, 133 studies were selected for full text screening. Of these, 24 studies were initially selected for inclusion; however, 7 studies were not included in the meta-analysis because the data overlapped with other included studies. Where data overlapped, the most relevant study was selected based on aims (i.e., studies where the primary aim addressed the question of interest were selected above those which addressed the question in secondary analysis) and sample size (i.e., larger sample sizes were included where both studies were relevant based on aims). In the meta-analysis, 16 studies were included in the main pooled unadjusted analysis and 17 studies were included in the pooled adjusted analysis. One study was excluded from the unadjusted analysis due to insufficient raw data availability but had adjusted results available.

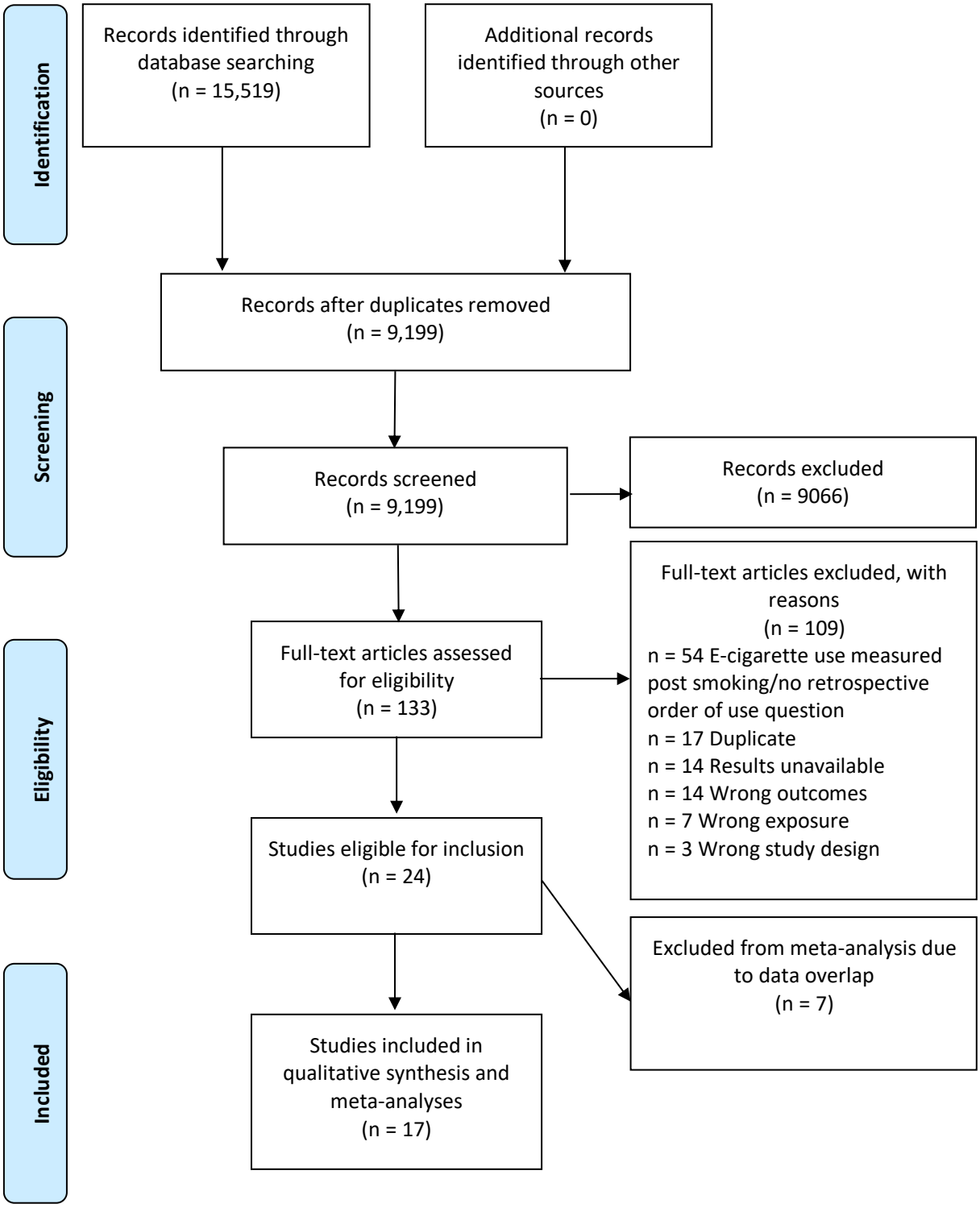


Figure 2.1. Study selection PRISMA flow diagram.

2.4.2 Study Characteristics

Full details of the study characteristics are shown in Appendix 3. All but one of the studies included were longitudinal. One study was cross-sectional in which participants were asked questions regarding their product use retrospectively (Auf et al., 2018). Total study sizes varied considerably, ranging from 347 to 39,718 and the number of participants included in the final analyses were often substantially smaller. Participants were mostly under 18 years of age and many of the studies were school-based. Where the numbers of males and females were reported, approximately 50% of the participants were male in most studies. However, only 33% were male in one study (Loukas, Marti, Cooper, Pasch, & Perry, 2018). Few studies reported the numbers of males and females by e-cigarette exposure, but for those that did, there were consistently a higher percentage of males in the exposed group compared with the unexposed group (Leventhal et al., 2015; Primack et al., 2018; Primack et al., 2015). The majority of the studies (10 out of 17) were conducted in the US, three studies took place in the UK (Best et al., 2018; Conner et al., 2018; East et al., 2018), one was based in Canada (Hammond, Reid, Cole, & Leatherdale, 2017), one in Mexico (Lozano et al., 2017), one in Germany (Morgenstern, Nies, Goecke, & Hanewinkel, 2018) and one in the Netherlands (Treur, Rozema, Mathijssen, van Oers, & Vink, 2018). Follow up periods in the longitudinal studies ranged from 4 to 24 months.

In terms of observed exposures, most studies explored ever e-cigarette use with never e-cigarette users as a comparator. Two studies looked at current e-cigarette users with not current users as the comparator (Hammond et al., 2017; Miech et al., 2017) and two looked at both current and ever e-cigarette use (Spindle et al., 2017; Watkins, Glantz, & Chaffee, 2018). Only one study considered the amount of exposure to nicotine (Treur et al., 2018), and one study looked at frequency of e-cigarette use (Wills et al., 2017). In the one study that took nicotine use into account (Treur et al., 2018), two separate analyses were conducted for: 1) ever use of nicotine containing e-cigarettes (OR = 11.90, 95% CI 3.36 to 42.11); and 2) ever use of non-nicotine containing e-cigarettes (OR= 5.36, 95% CI 2.73 to 10.52). However, the analysis groups were not mutually exclusive (i.e., an individual would have been in both analysis groups if they had tried both nicotine containing and nicotine free e-cigarettes). No analysis was reported using subgroups of exclusive nicotine or nicotine-free use. The one study which addressed frequency of

e-cigarette use (Wills et al., 2017) found that those who had used e-cigarettes at varying frequencies from once or twice (OR = 2.88, 95% CI 1.96 to 4.22) to weekly/daily (OR = 4.09, 95% CI 2.43 to 6.88) were more likely than those who had not used e-cigarettes to have smoked at least once at follow up.

Most of the included studies used ever smoking as an outcome. One study explored experimentation with smoking, as well as frequent and infrequent smoking (Barrington-Trimis et al., 2018), and three looked at recent/current smoking at follow up (Leventhal et al., 2015; Spindle et al., 2017; Watkins et al., 2018).

There was considerable variation between the covariates included in each study analysis (Appendix 3). One study only adjusted for four covariates (Best et al., 2018) while another adjusted for over 20 covariates (Morgenstern et al., 2018). All studies adjusted for sex and most adjusted for age and race/ethnicity. Other frequently included covariates were peer smoking, sensation seeking (and related factors), and drug and alcohol use.

2.4.3 Quality/Risk of Bias Within Studies and Causality

The quality of studies (or inversely, risk of bias) was good in most cases when rated using the NOS (Table 2.1). One study was rated as fair quality (Lozano et al., 2017), and three were rated as poor quality (Auf et al., 2018; Primack et al., 2018; Treur et al., 2018). Of the four Bradford-Hill criteria for causality deemed relevant to this research, the majority (11 studies) met three criteria (usually strength of evidence, temporality and specificity), four studies only met two criteria (Auf et al., 2018; Hammond et al., 2017; Leventhal et al., 2015; Loukas et al., 2018) and two met four criteria (Treur et al., 2018; Wills et al., 2017).

Table 2.1. Within study risk of bias and relevant Bradford Hill criteria for causality.

Study	Newcastle-Ottawa Scale Quality Rating	**Strength of Association** Strength of adjusted odds ratio (OR)	**Specificity** Adjusted for more than basic demographics	**Temporality** Longitudinally assessed	**Dose Responsivity** Frequency/ length of use/nicotine content taken into account	Number of Bradford-Hill criteria met out of 4
Auf et al. (2018)	Poor	Strong	Yes	No	No	2
Barrington-Trimis et al. (2018)	Good	Strong	Yes	Yes	No	3
Best et al. (2018)	Good	Strong	Yes	Yes	No	3
Conner et al. (2018)	Good	Strong	Yes	Yes	No	3
East et al. (2018)	Good	Strong	Yes	Yes	No	3
Hammond et al. (2017)	Good	Strong	No	Yes	No	2
Leventhal et al. (2015)	Good	Weak	Yes	Yes	No	2
Loukas et al. (2018)	Good	Weak	Yes	Yes	No	2
Lozano et al. (2017)	Fair	Strong	Yes	Yes	No	3
Miech et al. (2017)	Good	Strong	Yes	Yes	No	3
Morgenstern et al. (2018)	Good	Strong	Yes	Yes	No	3
Primack et al. (2015)	Good	Strong	Yes	Yes	No	3
Primack et al. (2018)	Poor	Strong	Yes	Yes	No	3
Spindle et al. (2017)	Good	Strong	Yes	Yes	No	3
Treur et al. (2018)	Poor	Strong	Yes	Yes	Yes	4
Watkins et al. (2018)	Good	Strong	Yes	Yes	No	3
Wills et al. (2017)	Good	Strong	Yes	Yes	Yes	4

Note: Thresholds were applied to convert the Newcastle-Ottawa scales to Agency for Health Research and Quality standards (whereby a good quality rating indicates low risk of bias and a poor rating indicates high risk of bias). Good quality = 3/4 stars in selection domain AND 1/2 stars in comparability domain AND 2/3 stars in outcome/exposure domain. Fair quality = 2 stars in selection domain AND 1/2 stars in comparability domain AND 2/3 stars in outcome/exposure domain. Poor quality = 0/1 star in selection domain OR 0 stars in comparability domain OR 0/1 stars in outcome/exposure domain. ORs were described as strong if more than 2 and weak if less than or equal to 2.

2.4.4 Results of Individual Studies

The results of individual studies included in the main meta-analysis can be found in Table 2.2 and within forest plots in Figures 2.2 (unadjusted) and 2.3 (adjusted). Effect sizes (ORs) ranged from 2.46 to 12.31 (unadjusted). All estimates were considered to show strong evidence of a positive association between e-cigarette use among non-smokers and later smoking in unadjusted analyses. Covariates included in the adjusted analyses varied on a study-by-study basis. After adjustment, effect estimates in all but three studies (Leventhal et al., 2015; Loukas et al., 2018; Lozano et al., 2017) remained strong (i.e., OR > 2). These studies included relatively rigorous adjustment for covariates (i.e., adjusting for peer use, impulsivity, susceptibility etc.). The inclusion of covariates in the model attenuated most results towards the null (although none crossed the null). Effect sizes were strengthened after adjustment in four studies (Barrington-Trimis et al., 2018; Primack et al., 2018; Primack et al., 2015; Treur et al., 2018).

Table 2.2. Individual results of studies included in the meta-analysis.

	Initiated cigarette smoking (n)/ not initiated cigarette smoking (n)		Odds of initiating smoking	
Study	E-cigarette users	Never/not current e-cigarette users	Unadjusted odds ratio (95% CI)	Adjusted odds ratio (95% CI)
Auf et al. (2018)	*	*	*	3.7 (3.1, 4.5)*
Barrington-Trimis et al. (2018)	184/857	280/4,171	3.80 (3.10, 4.66)	4.57 (3.56, 5.87)
Best et al. (2018)	74/183	249/1,942	4.62 (3.34, 6.38)	2.42 (1.63, 3.60)
Conner et al. (2018)	118/343	124/1,383	5.32 (3.99, 7.11)	4.06 (2.94, 5.60)
East et al. (2018)	11/21	74/902	12.31 (5.06, 29.94)	10.57 (3.33, 33.50)
Hammond et al. (2017)	136/487	1,313/16,831	4.58 (3.73, 5.63)	2.12 (1.68, 2.66)
Leventhal et al. (2015)	19/222	71/2,308	2.95 (1.74, 4.99)	1.75 (1.10, 2.77)
Loukas et al. (2018)	114/568	168/1,190	2.72 (2.10, 3.53)	1.36 (1.01, 1.83)
Lozano et al. (2017)	86/216	950/4,479	2.46 (1.85, 3.26)	1.60 (1.31, 1.97)**
Miech et al. (2017)	4/13	14/213	6.32 (1.73, 23.10)	6.58 (2.04, 57.88)**
Morgenstern et al. (2018)	93/312	175/1,867	4.11 (3.08, 5.48)	2.50 (1.82, 3.54)**
Primack et al. (2015)	6/16	65/678	5.66 (1.99, 16.07)	8.30 (1.20, 58.60)
Primack et al. (2018)	6/16	81/889	6.06 (2.15, 17.10)	6.82 (1.65, 28.25)
Spindle et al. (2017)	45/153	230/2,163	3.50 (2.41, 5.09)	3.37 (1.91, 5.94)
Treur et al. (2018)	432/740	235/2,049	10.83 (8.87, 13.22)	11.9 (3.36, 42.11)
Watkins et al. (2018)	81/425	387/9,923	5.80 (4.46, 7.54)	2.53 (1.80, 3.56)
Wills et al. (2017)	42/215	50/926	4.25 (2.74, 6.61)	2.87 (2.03, 4.05)
Overall	1,451/4,787	4,340/52,727	4.59 (3.60, 5.85)	2.92 (2.30, 3.71)

*Raw data were not available or insufficient information was provided to calculate an accurate unadjusted odds ratio. Adjusted results were reported to one decimal place. **Estimates were provided as risk ratios and converted to odds ratios.

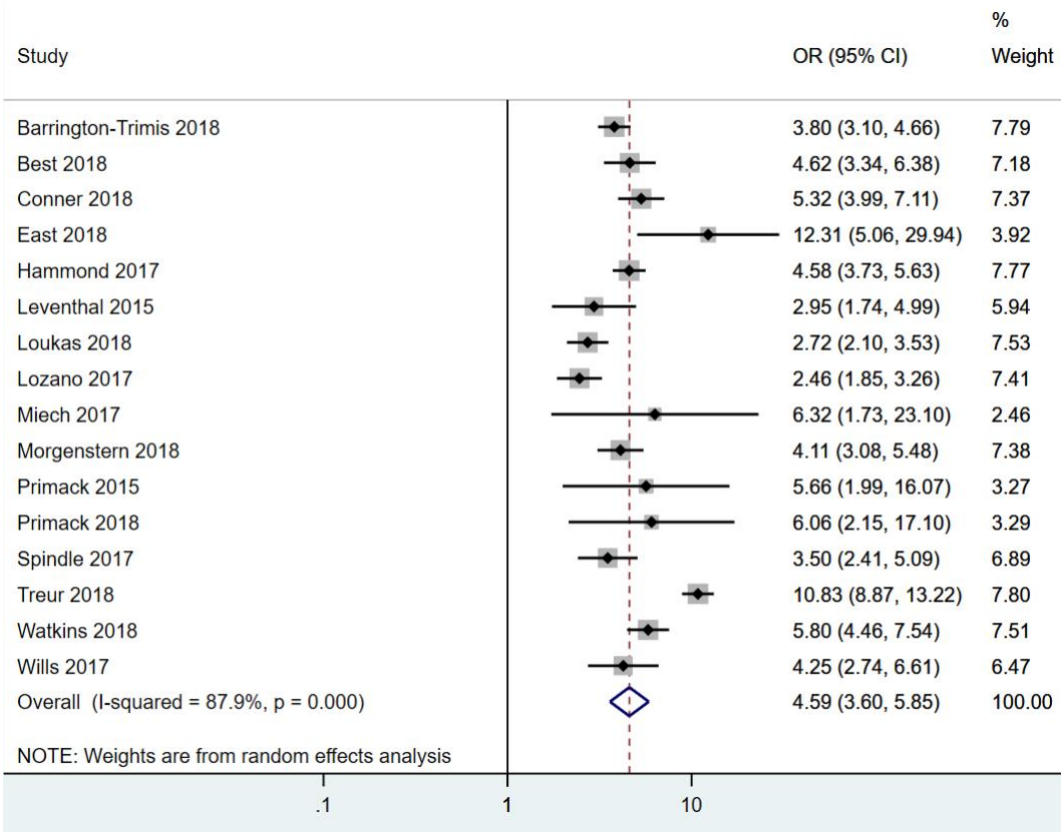


Figure 2.2. Forest plot for the unadjusted association between e-cigarette use and subsequent smoking.

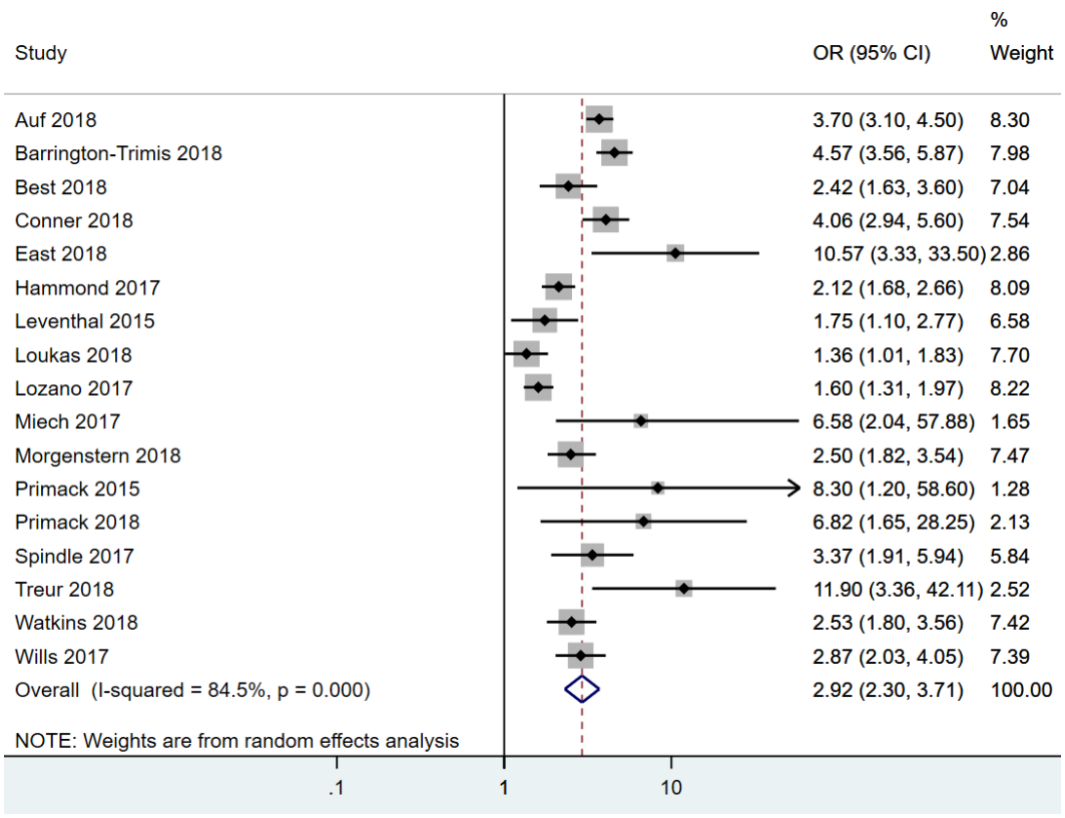


Figure 2.3. Forest plot for the adjusted association between e-cigarette use and subsequent smoking.

2.4.5 Synthesis of Results

When pooled in a random effects meta-analysis, e-cigarette use in non-smoking young people was associated with a 4-and-a-half-fold increase in the odds of subsequent smoking (unadjusted; OR = 4.59, 95% CI 3.60 to 5.85). Pooling the adjusted estimates, the association was still strong but somewhat weaker (adjusted; OR = 2.92, 95% CI 2.30 to 3.71). Heterogeneity statistics indicated there was high heterogeneity in both the unadjusted (I^2 = 88%) and adjusted (I^2 = 85%) analyses.

Of the 16 studies included in the unadjusted meta-analysis, 13 provided results which explored ever e-cigarette use and ever smoking and three studies provided results for ever e-cigarette use and current smoking. Three studies explored past 30-day e-cigarette use and ever smoking. Further studies explored: past 30-day use of e-cigarettes and past 30-day smoking (one study), frequency of e-cigarette use (one study) and frequency of smoking (one study). Pooled analyses require at least two studies to be available, thus pooled analyses were not possible for these subgroups. Forest plots of the analyses sub-grouped by varying exposure and outcome definitions (adjusted estimates) can be found in Figures 2.4-2.7. The number of studies included in the adjusted analyses were slightly higher in some cases than the unadjusted results due to the availability of raw data to calculate an unadjusted estimate (Auf et al., 2018; Spindle et al., 2017).

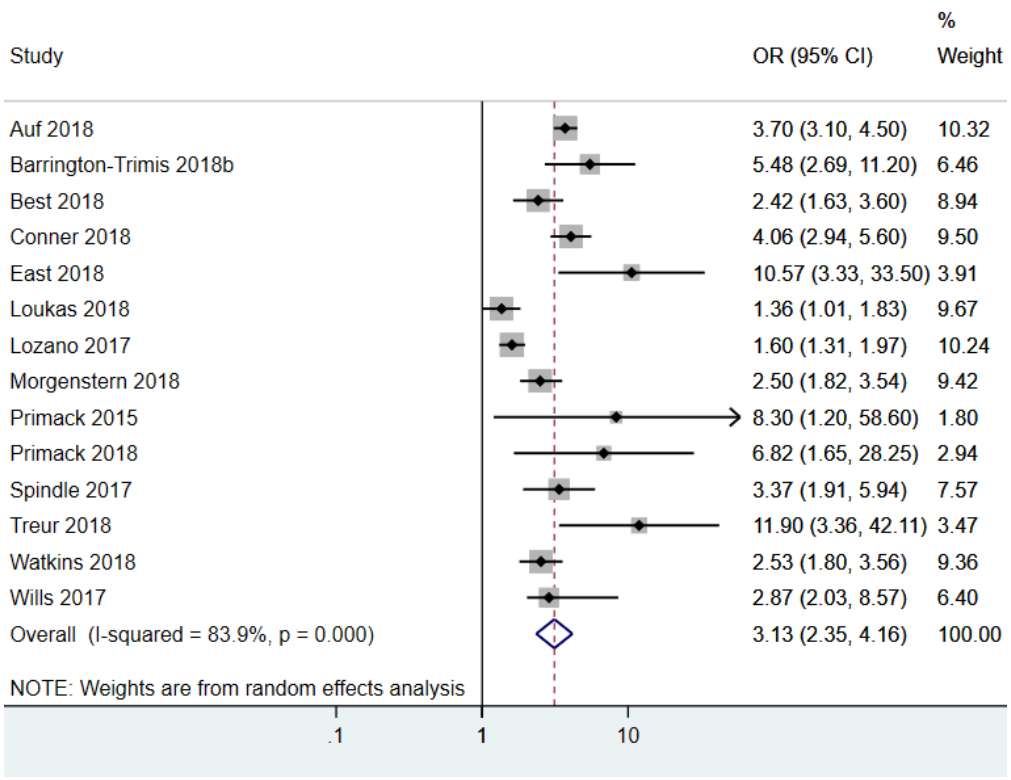


Figure 2.4. Forest plot for the adjusted association between ever e-cigarette use and later ever smoking.

2.4.5.1 Ever e-cigarette use and ever smoking (subgroup 1). Pooled analyses of studies exploring ever vaping among never smokers and subsequent ever smoking resulted in a pooled unadjusted OR of 4.17 (95% CI 3.53 to 6.29). Heterogeneity between included studies in this analysis was high (I^2 = 90%). The results of the pooled adjusted analysis (Figure 2.4) were similar with a slightly lower odds ratio (OR = 3.13, 95% CI 2.35 to 4.16). Heterogeneity between included studies was still high in the adjusted analysis (I^2 = 84%).

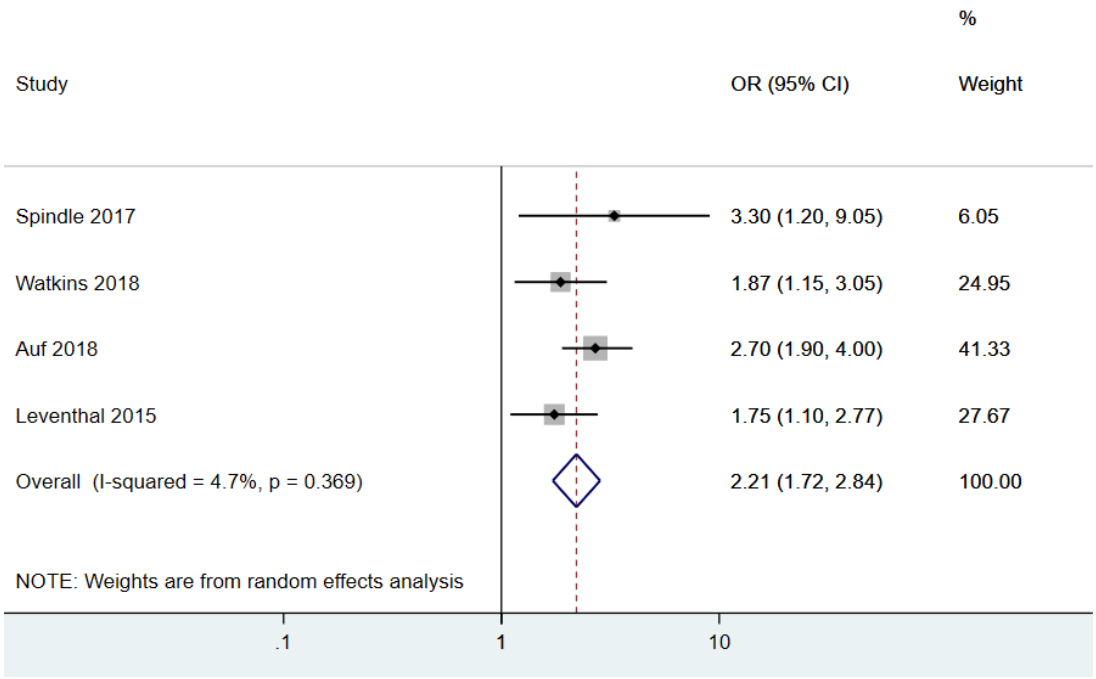


Figure 2.5. Forest plot for the adjusted association between ever e-cigarette use and later current smoking.

2.4.5.2 Ever e-cigarette use and current smoking (subgroup 2). Among never smokers, pooled unadjusted analyses indicated ever e-cigarette users had increased odds of subsequently becoming a current smoker (OR = 4.35, 95% CI 2.95 to 6.42) compared with never e-cigarette users. Heterogeneity estimates indicated low heterogeneity (I^2 = 41%). Adjusted analyses (Figure 2.5) showed similar but weakened results when pooled (OR = 2.21, 95% CI 1.72 to 2.84). Heterogeneity was indicated as low in the adjusted pooled analyses (I^2 = 5%).

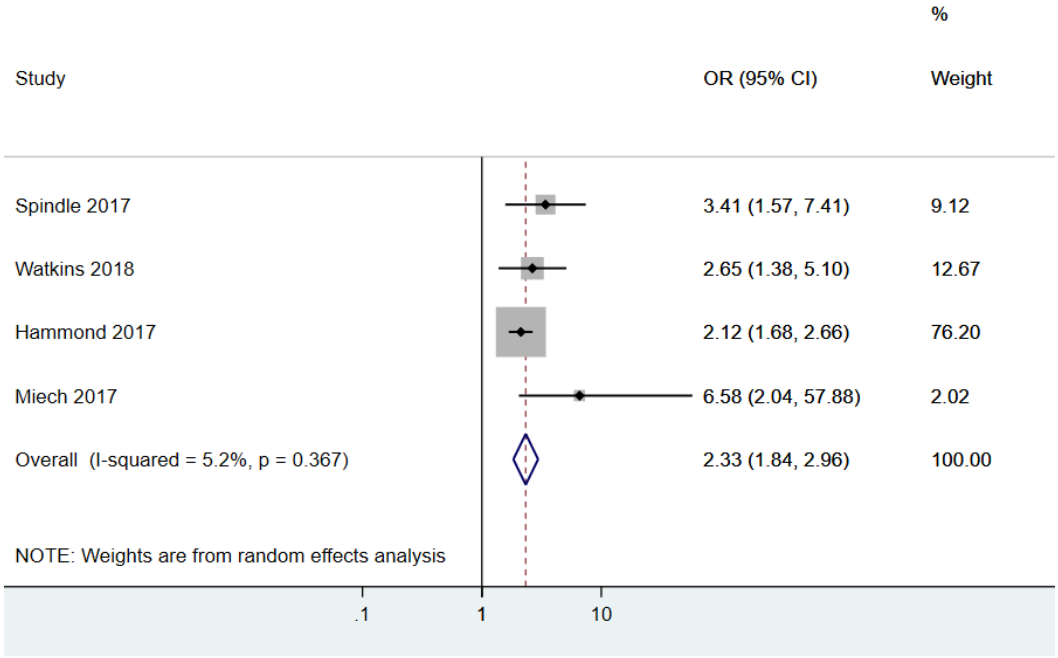


Figure 2.6. Forest plot for the adjusted association between past 30-day e-cigarette use and later ever smoking.

2.4.5.3 Current e-cigarette use and ever smoking (subgroup 3). Past 30-day use of e-cigarettes among never smokers was associated with increased odds of ever subsequently smoking (OR = 5.64, 95% CI 3.75 to 8.50) in pooled unadjusted analysis. Heterogeneity estimates indicated moderate heterogeneity (I^2 = 49%). The pooled adjusted analysis (Figure 2.6) also indicated increased odds of ever subsequently smoking (OR = 2.33, 95% CI 1.84 to 2.96). Heterogeneity estimates also indicated low heterogeneity when the adjusted results were pooled (I^2 = 5%).

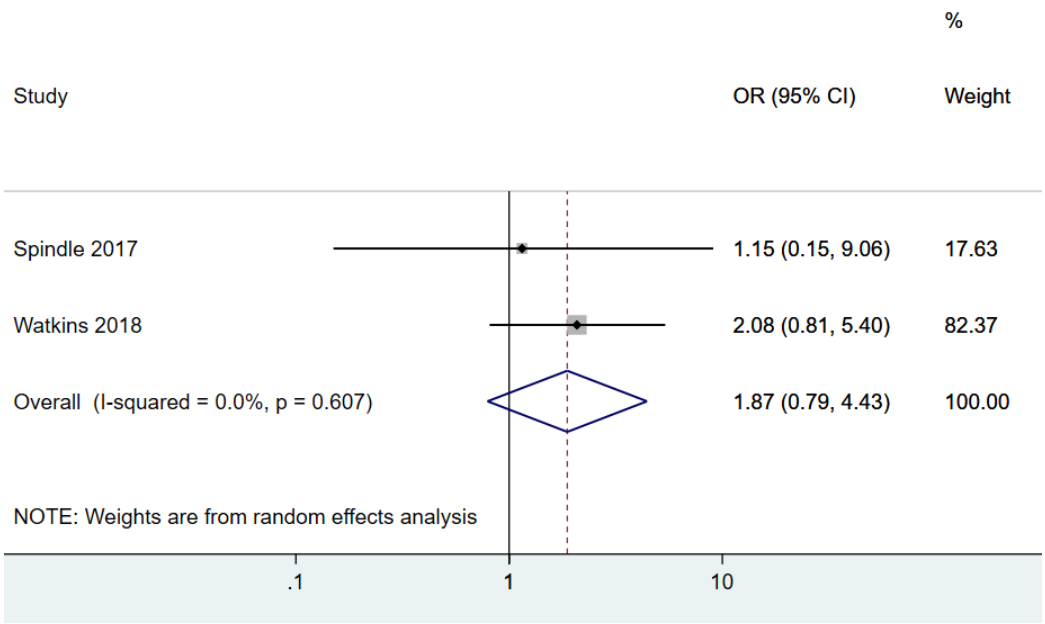


Figure 2.7. Forest plot for the adjusted association between current e-cigarette use and later current smoking.

2.4.5.4 Current e-cigarette use and current smoking (subgroup 4). There was insufficient data available to calculate unadjusted estimates for one of the two studies which reported adjusted estimates for past 30-day use of e-cigarette and past 30-day smoking. Consequently, I was only able to pool the adjusted estimates. There was no clear evidence of an association between current e-cigarette use and current smoking at follow up in the pooled adjusted analyses (Figure 2.7).

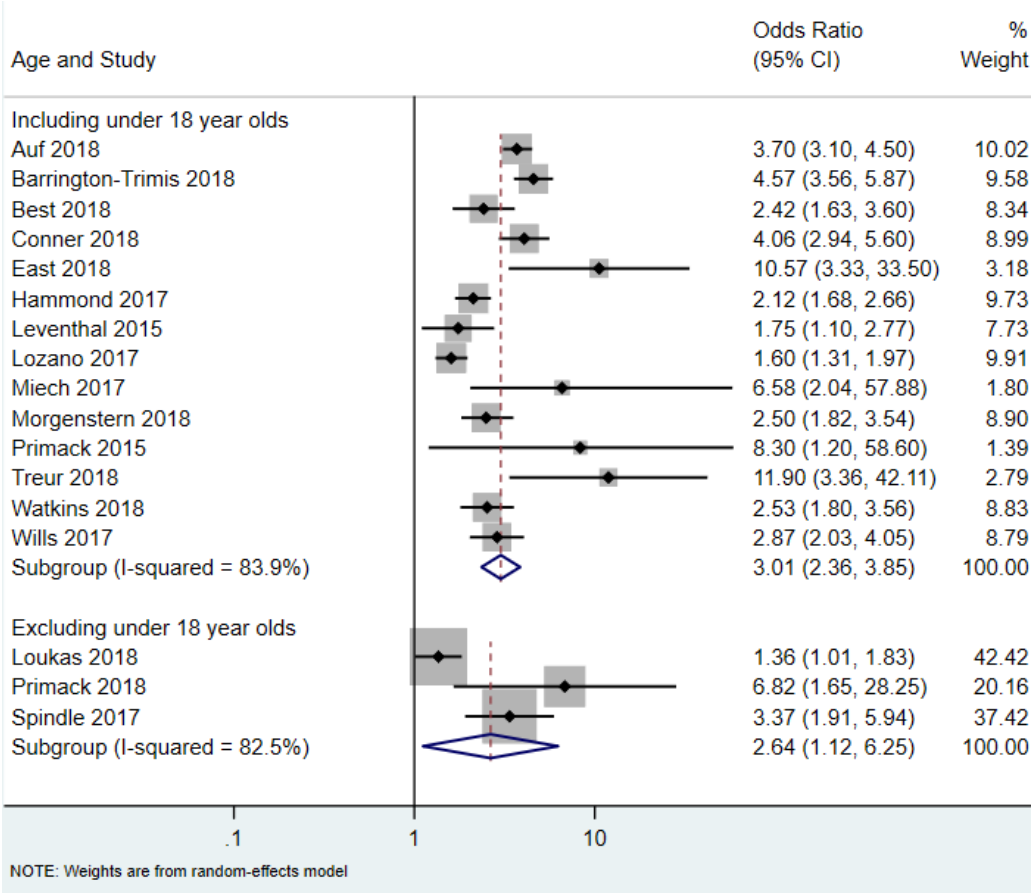


Figure 2.8. Forest plot for the adjusted association between e-cigarette use and later smoking among studies including and excluding those under the age of 18 years.

2.4.5.5 Stratification by age. When the unadjusted main analyses were stratified by age (including versus excluding those under the age of 18 year olds) the pooled OR among studies including those under the age of 18 years was slightly higher (OR = 4.87, 95% CI 3.73 to 6.35) than the pooled OR of studies excluding those under 18 (OR = 3.17, 95% CI 2.37 to 4.25). Heterogeneity estimates indicated that there was low heterogeneity between studies excluding under 18's (I^2 = 32%) but high heterogeneity between studies including under 18's (I^2 = 88%). Adjusted pooled analyses including and excluding those under the age of 18 years are reported in Figure 2.8. The pattern of the results is similar in the adjusted and unadjusted analysis, but there was more evidence of heterogeneity in the adjusted analysis excluding those under 18 than in the unadjusted analysis. Of note, the odds of smoking reduced by half after adjustment for one study (Loukas et al., 2018).

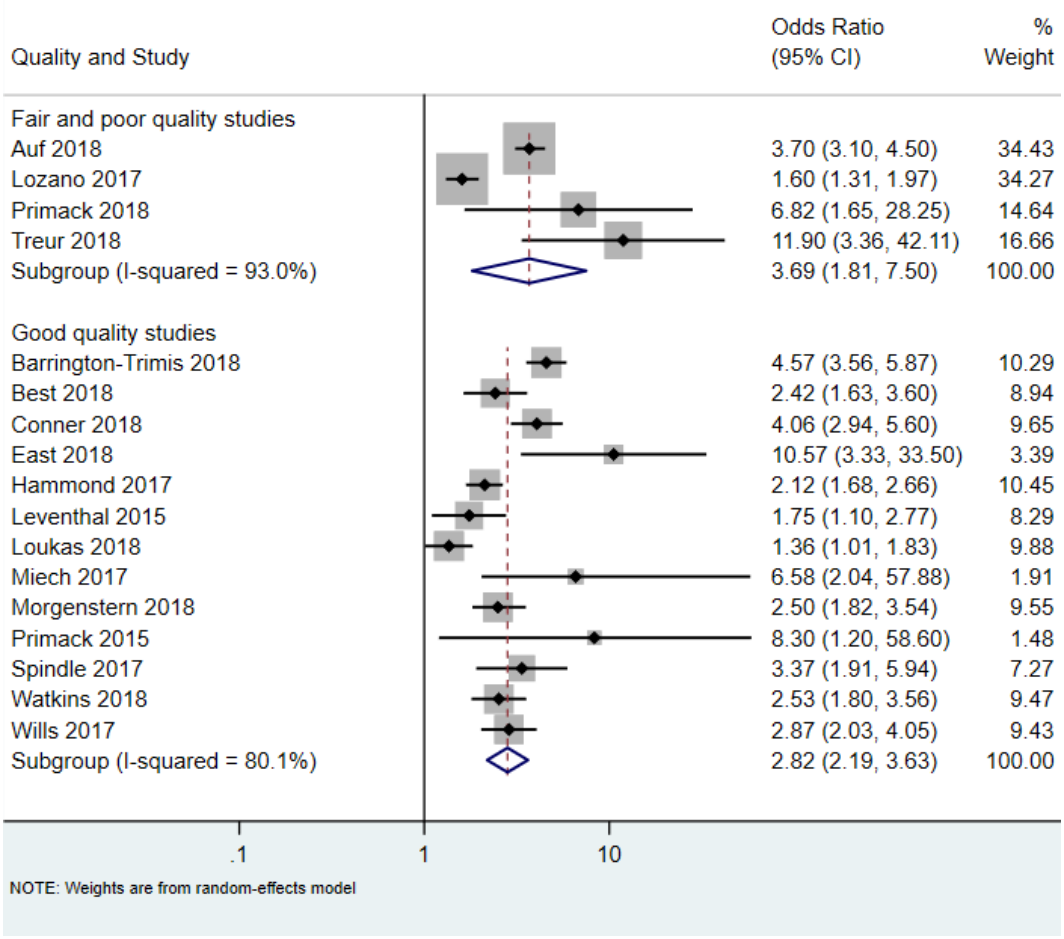


Figure 2.9. Forest plot for the adjusted association between e-cigarette use and later smoking among studies of fair/poor quality and good quality.

2.4.5.6 Stratification by study quality/risk of bias. Due to limited variation in quality rating of studies using the NOS, fair and poor quality studies were pooled and compared with good quality studies. The pooled unadjusted OR for studies rated as good quality (n = 13 studies; OR = 4.29, 95% CI 3.67 to 5.01) was lower than the pooled ORs for fair/poor quality studies (n = 3 studies; OR = 5.41, 95% CI 1.67 to 17.51), but the confidence intervals overlap. Heterogeneity measures indicated that high quality studies were less heterogeneous than fair/poor quality studies (I^2 = 60% and I^2 = 97% respectively). Adjusted pooled analyses are shown in Figure 2.9. The pattern of the results is similar in the adjusted and unadjusted analysis, but there was more evidence of heterogeneity after adjustment, particularly for good quality studies.

2.4.5.7 Stratification by support for the gateway hypothesis. During the review process it became apparent that many studies did not draw clear conclusions regarding the gateway hypothesis or made balanced conclusions. This made it difficult to categorise studies and as such I was unable to stratify based on this criterion. This is encouraging as it suggests that authors are somewhat acknowledging the limitations of their findings, however, it could equally be a result of authors avoiding discussing the gateway hypothesis as it is such a charged topic. In hindsight, clearer criterion for determining a study to be supporting or refuting the hypothesis should have been outlined in my study protocol to allow my review team and I to more effectively judge the level of support for the gateway hypothesis that was present in each study conclusion. However, the use of strict criteria could distort the strength of the authors' conclusions and inflate the appearance of bias by ignoring the nuance within the conclusion.

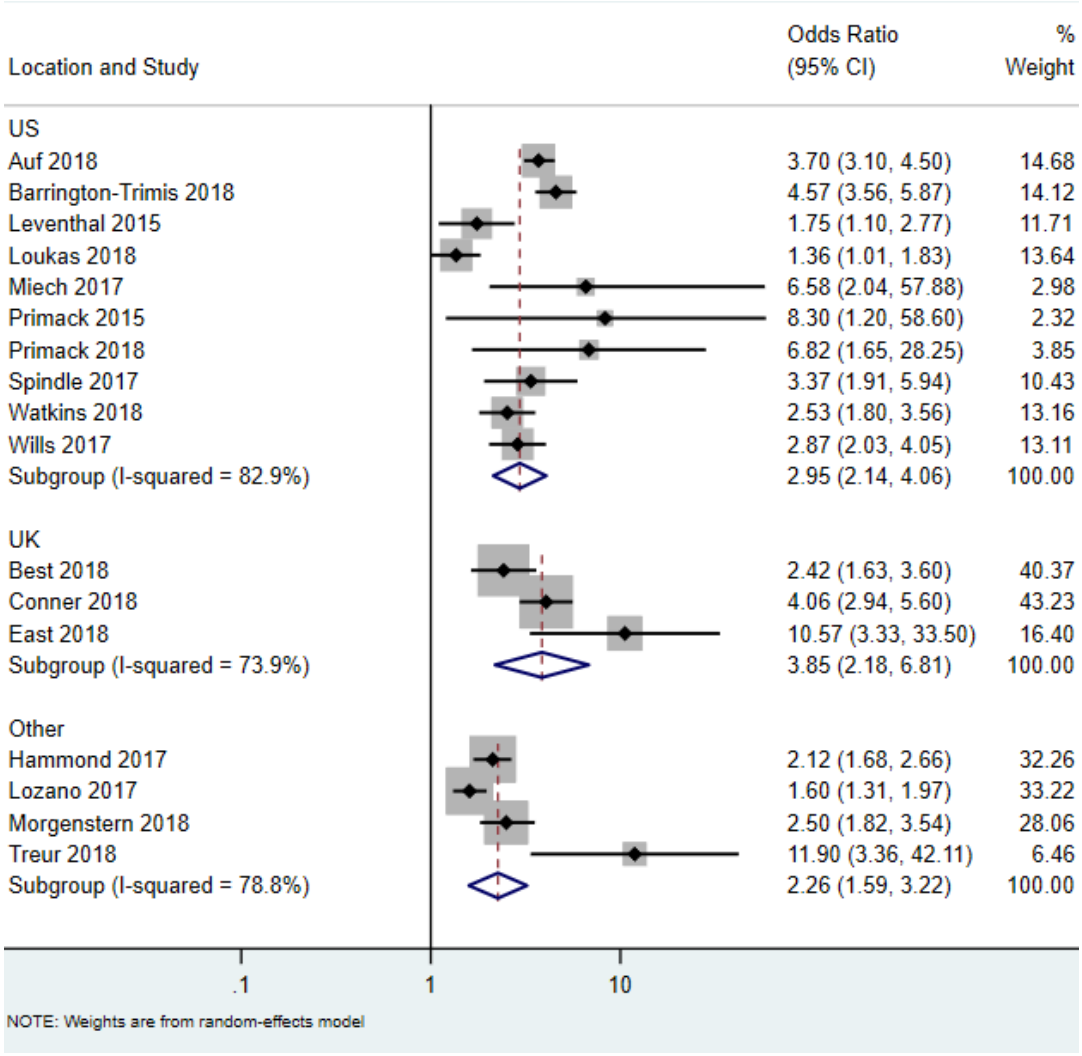


Figure 2.10. Forest plot for the adjusted association between e-cigarette use and later smoking among US studies, UK studies, and studies outside of the US and UK.

2.4.5.8 Stratification by location of study. There were only two countries in which more than one included study was conducted; 10 studies took place in the US and three took place in the UK. The pooled estimate for unadjusted odds of studies conducted in the US was 3.95 (95% CI 3.17 to 4.92) and in the UK was 5.55 (95% CI 3.94 to 7.82). Heterogeneity for studies conducted was high in the US (I^2 = 93%) and moderate in the UK (I^2 = 52%). The remaining studies were located in the Netherlands, Germany, Canada and Mexico. Pooling the results of these four studies, the odds of subsequent smoking was 4.75 (95% CI 2.54 to 8.89) with high heterogeneity between studies (I^2 = 96%). Adjusted pooled analyses are shown in Figure 2.10.

2.4.6 Risk of Bias Across Studies

The risk of bias across studies is shown in a funnel plot in Figure 2.11. The figure is somewhat asymmetrical and some points (25%) lie above the superimposed funnel limits (95% confidence region) suggesting that there may be some publication bias and indicating there may be heterogeneity (as supported by the I^2 statistics) or selection bias across the included studies (Sterne & Harbord, 2004).

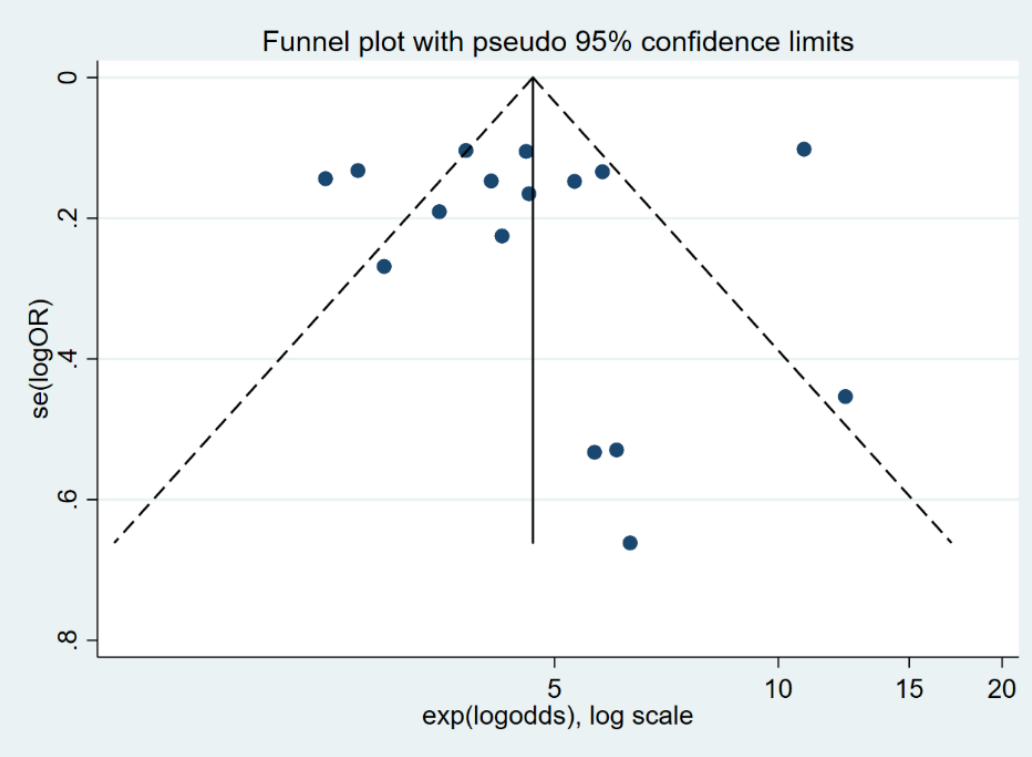


Figure 2.11. Funnel plot to assess risk of bias across studies.

2.5 Discussion

I have provided an update to the previous meta-analysis by identifying an additional 11 studies which were not included in Soneji and colleagues' (2017) meta-analysis and 6 studies which were not included in the systematic review by Glasser and colleagues (2018). One study which was included in the previous meta-analysis was not included in this meta-analysis; it was substituted for more recent evidence using an overlapping data set (Barrington-Trimis et al., 2016). In addition, one conference abstract which is no

longer accessible was substituted for the published article (Primack et al., 2016). The pooled adjusted estimate indicated a strong association, reflecting a nearly 3-fold increased odds of later smoking among young people who have used e-cigarettes. Subgrouping and stratification revealed some differences between groups, but consistently indicated a strong positive association. The main findings were consistent with the findings of Soneji and colleagues (2017), whereby there was a strong positive association between e-cigarette use among non-smokers and subsequent smoking and a high degree of heterogeneity between studies. Similar to Soneji and colleagues (2017), stratification by age revealed slightly lower pooled estimates for the odds of smoking in studies which excluded those under the age of 18 years compared with studies including them. Stratification by location indicated stronger associations in the UK compared with the US and other countries.

Three of the four pre-selected Bradford-Hill criteria for causality (strength of association, temporality and specificity) were commonly rated as having been met in the included studies. Although not assessed on a study-by-study basis, the consistency and plausibility criteria were also met; the included study estimates are *consistently* in the same direction and *plausible* causal pathways were considered or discussed in some studies (e.g., nicotine addiction or similar hand-to-mouth actions for both behaviours). Although meeting these criteria provides some support for a possible causal relationship between e-cigarette use and later smoking, which is in line with the theory that e-cigarettes act as a gateway to smoking (Bell & Keane, 2014; Etter, 2018), meeting these criteria does not rule out other potential explanations of the association (e.g., a genetic or common liability).

Commonly in the literature, the claim is that the gateway effect is attributable to nicotine addiction (Bell & Keane, 2014). E-cigarettes have historically not delivered nicotine as effectively as cigarettes (Farsalinos et al., 2014), so that e-cigarettes may not be adequate to satisfy users who become more heavily addicted to nicotine (Vanyukov et al., 2012). In contrast, the common liability theory proposes that people who use multiple drugs (or in this case different delivery methods of the same drug) share the same predisposing factors (Chapman et al., 2019). When considering a causal relationship between e-cigarette use and later smoking (e.g., the gateway hypothesis), these shared factors/confounders (i.e., the common liability theory) should be taken into account. Due to a lack of consideration of such factors in the included studies, I

have strong reservations about inferring a causal relationship from my findings. I discuss my reservations below with respect to the Bradford-Hill criteria for causality, while highlighting the additional research which may help to strengthen the evidence for a causal relationship. I have also described how identifying areas for improvement in the field helped to inform my research questions and study designs in the subsequent chapters of this thesis.

The majority of studies included in the meta-analysis satisfied the temporality criterion. All studies except one were longitudinal and measured exposure prior to smoking and smoking at follow up; therefore, the observed association between earlier e-cigarette use and later smoking is unlikely to be due to reverse causality (i.e., smoking leading to e-cigarette use). However, smoking status is sometimes misreported by young people (Khouja, Munafò, et al., 2020) meaning that some self-reported non-smokers at baseline may in fact have smoked previously. If ever smokers are more likely to use e-cigarettes and are more likely to misreport their smoking history at baseline compared with never smokers, but accurately report that they are ever smokers at follow up, the association could be biased away from the null. Self-reports were validated in one study (Conner et al., 2018) using breath carbon monoxide levels; however, with the short half-life of breath carbon monoxide (4-6 hours) this measure is not suitable to validate self-reports of ever smoking. Other methods of biochemical verification of smoking, such as assessments of cotinine levels, are unsuitable for validation in this research; cotinine is a metabolite of nicotine and will therefore be affected by both smoking and using nicotine-containing e-cigarettes, so it cannot be used to distinguish between products. In future research, the use of biomarkers which can identify long-term smoke exposure could help researchers to objectively confirm self-reported smoking status at baseline. Differential DNA methylation signals have been observed among smokers compared with non-smokers (Lee et al., 2015; Zeilinger et al., 2013) and it is plausible that different signals would be observed among e-cigarette users (Richmond et al., 2018). Eventually, researchers may be able to use these differential methylation signatures to exclude people who misreport their behaviour. Consequently, researchers could more confidently dismiss reverse causation as an explanation for the association.

It is also worth noting that the measure I used to determine specificity was relatively liberal. The review team and I rated studies as specific if they adjusted for more than basic demographic factors. In a cross-sectional study, accounting for shared risk factors

fully explained the relationship between e-cigarette use and current smoking, demonstrating the importance of adjusting for potential confounders such as alcohol and drug use, peer smoking and risk-taking behaviour (Kim & Selya, 2019). In addition, some potential confounders, like impulsivity, are difficult to fully capture via self-report and are often assessed relatively crudely. Had we considered only studies adjusting for behavioural risk factors (e.g., alcohol and marijuana use) to meet this criterion, only six studies would have been rated as specific. Adjusting for these factors reduced the OR and/or widened the confidence interval compared with the unadjusted analysis (ORs reduced by 1.40 on average), however, I found no clear difference in OR reduction when compared with studies not adjusting for behavioural risk factors (p = 0.19). Further research is needed to investigate the demographic, behavioural, and lifestyle factors which are associated with the use of e-cigarettes among specific populations. In Chapter 3, I have explored these factors among young adults in the UK.

Since statistical adjustment can never fully remove the risk of confounding, other approaches to exploring the potential for a common liability (e.g., to risk-taking) explaining the observed association between e-cigarette use and smoking are also warranted. One way of exploring the specificity of this relationship more thoroughly (which none of the included studies did) would be the use of negative control outcomes (i.e., outcomes which have similar confounding structures to smoking but for which there is no biologically plausible mechanism for e-cigarette use being a causal factor). For example, behaviours like smoking or using an e-cigarette are unlikely to cause other risk-taking behaviour such as increasing the number of sexual partners a person has; if similar associations are seen between e-cigarette use, smoking and number of sexual partners, it would indicate that the link may be caused by common underlying factors. Furthermore, exploring the genetic aetiology of e-cigarette use may help in understanding whether e-cigarette use and smoking share a common liability. If e-cigarette use has a shared genetic aetiology with negative control outcomes for which there are no plausible pathways through which e-cigarette use would be a causal factor (e.g., gambling) this would suggest that the association is due to the two behaviours sharing a common genetic liability for risk-taking. I have begun to explore this in Chapter 4 using polygenic risk scores of smoking initiation and utilising negative control outcomes, but the triangulation of evidence obtained using different methods will be critical here (Munafo & Davey Smith, 2018).

To meet the final pre-selected Bradford-Hill criterion, dose-response, increased e-cigarette use should lead to greater risk of later smoking. Despite the most likely causal pathway from e-cigarette use to later smoking being via nicotine addiction (Etter, 2018), only one of the included studies measured and took into account the nicotine content of the e-cigarettes used. This study indicated that both use of nicotine containing e-cigarettes and (to a lesser extent) non-nicotine containing e-cigarettes are strongly associated with later smoking (Treur et al., 2018). This suggests that nicotine exposure may be one factor in the association between e-cigarette use and later smoking, but not the sole mechanism. Unfortunately, the study reported analyses based on nicotine versus non-nicotine vaping in which these two groups were not mutually exclusive (i.e., individuals would be in both analysis groups if they tried both nicotine containing and nicotine free e-cigarettes). Thus, it is unclear whether there is an association between e-cigarette use among non-smokers and later smoking when users have not been exposed to nicotine. To determine whether there is a nicotine dose-response involved in the association, we would also need to observe the frequency of e-cigarette use prior to smoking. The one study that looked at frequency of use (at 4 levels) indicated that there may be a dose-response to nicotine when comparing use just once/twice to use weekly/daily (Wills et al., 2017), but the odds of later smoking did not increase linearly with each increased level of frequency of use. As nicotine is metabolised differently on an individual basis, direct measures of nicotine rather than frequency of use may be necessary to determine whether there is a dose-response. Although nicotine may play a causal role in the relationship between e-cigarette use and smoking, without further study of any dose-response relationship (including the study of nicotine content and frequency of e-cigarette use), I cannot confidently infer causality according to Bradford-Hill criteria.

The heterogeneity statistics of 88% (unadjusted analyses) and 85% (adjusted analyses) indicate that my effect estimates should be interpreted with caution. When studies included in meta-analyses are substantially statistically heterogenous (i.e., they vary in population, study design, risk of bias etc.), estimates of the combined effect may not be meaningful and confidence in the generalisability of the findings is reduced (Higgins et al., 2008; Higgins, Thompson, Deeks, & Altman, 2003). Although there is no commonly agreed upon strict threshold for heterogeneity, as a rough guide the Cochrane handbook suggests that I^2 statistics of between 75% and 100% represent considerable

heterogeneity (Higgins et al., 2008). Age contributed to the observed heterogeneity in this meta-analysis – the association was stronger in studies including those under the age of 18 years than studies excluding them. In adolescence, risk-taking is common (Reyna & Farley, 2006) and decision making for health-risk behaviours is influenced by peers, societal influences and parental monitoring, but these factors are less influential to adults (Halpern-Felsher, Baker, & Stitzel, 2016). Such factors are likely to be confounders of the association between e-cigarette use and later smoking (particularly in studies of those under the age of 18 years) and should be included as covariates where possible. The results stratified by location also suggested there may be societal influences on the association; the association was stronger among studies based in the UK than those based in the US. This suggests that country-specific societal factors, such as legislation, taxation, social norms, and public opinion, may be confounding this association such that study results may not be generalisable to other countries.

Considering the reservations which I have outlined, it is still unclear whether e-cigarette use is causing young people to start smoking. However, a substantial number (10%) of young people were identified in these studies who have used e-cigarettes as non-smokers. Although many of the studies failed to account for whether the young people were using e-cigarettes regularly or frequently (rather than just experimenting), this does raise concerns about the potential health impact of long-term nicotine use aside from exposure to other constituents of cigarette smoke. Future studies should aim to understand why youth and young adults vape, and whether this relates to continued smoking and vaping and should employ novel methods to explore the potential health effects related to long-term nicotine use. For example, in this thesis I have explored the smoking and vaping behaviour associated with different reasons for vaping (Chapter 3) and I have used multivariable Mendelian randomisation (MVMR) to explore the effects of nicotine on various health outcomes while accounting for cigarette consumption in general, and vice versa (Chapter 5).

2.6 Chapter Summary

There is a strong consistent association in observational studies between e-cigarette use among non-smokers and later smoking. However, findings from published observational studies do not provide clear evidence that this is explained by a gateway effect rather than shared common causes of both use of e-cigarettes and smoking. Future research

should explore the factors influencing e-cigarette use and consider including relevant potential confounders, such as better measures of impulsivity and other measures of propensity to risk-taking, as well as objective measures of smoking status in order to better explore the potential role of e-cigarettes as a gateway to smoking. Studies that explore the genetic underpinnings of these behaviours and use negative control outcomes may also help improve our understanding of the association between e-cigarette use and later smoking. Finally, to prevent the use of e-cigarettes among young people (and potential consequences of e-cigarette use, such as smoking), it is important to identify adolescents and young adults who are more likely to use e-cigarettes. It is also necessary to understand why they use e-cigarettes and explore whether this impacts their longer-term smoking and vaping behaviour. In the next Chapter, I start to explore this by i) looking at the characteristics of e-cigarette users and never users to identify which young adults are more likely to use e-cigarettes; and ii) exploring the association between different reasons for vaping and later smoking and vaping status to identify subgroups of users who are more likely to continue to smoke and/or vape.

Chapter 3 Who uses and continues to use e-cigarettes and why?

This chapter closely resembles sections from the following pre-print:

Khouja, J. N., Taylor, A. E., & Munafò, M. R. (2020). Associations between reasons for vaping and current vaping and smoking status: Evidence from a UK based cohort. *Drug and Alcohol Dependence, 217,* 108362; doi.org/10.1016/j.drugalcdep.2020.108362.

Khouja, J. N., Taylor, A. E., & Munafò, M. R. (2020). Associations between reasons for vaping and current vaping and smoking status: Evidence from a UK based cohort. *medRxiv***;** doi:10.1101/19006007**.**

I took the lead in designing, implementing and interpreting this observational study with the support of my co-authors. Specifically, MM and AT provided feedback on the design of the study and during the preparation of the manuscript. I was responsible for the data analysis and was responsible for writing the first draft of the manuscript and editing the manuscript in response to comments from co-authors and reviewers.

3.1 Chapter Overview

As discussed in Chapters 1 and 2, it is important to know who is vaping, why, and how this may relate to their later smoking and vaping behaviour. In this Chapter, I compare the characteristics of ever vapers versus never vapers in a cohort of 3,994 young adults in the Avon Longitudinal Study of Parents and Children (ALSPAC). I further explore the characteristics of ever vapers by comparing former and current vapers. For ever vapers, I describe the retrospectively self-reported reasons for vaping by 23 years of age and explore the association with vaping and smoking status at 24 years. Using logistic regression, I assess the association with subsequent vaping behaviour among ever vapers who had ever smoked, and with subsequent smoking behaviour among individuals who had been regular smokers prior to vaping. I restrict my analysis to ever smokers due to limited numbers of never smokers who had vaped by 23 years.

3.2 Introduction

Of an estimated 3.6 million vapers in Great Britain, over half now consider themselves ex-smokers (Action on Smoking and Health, 2019a). As discussed in Chapter 1, evidence suggests e-cigarettes are less harmful than cigarettes (Public Health England, 2015) and can be an effective smoking cessation aid (Hartmann-Boyce et al., 2020). Nevertheless, 34% of adult smokers have not tried e-cigarettes (Action on Smoking and Health, 2019a), and not all who have tried them have successfully quit smoking (Hartmann-Boyce et al., 2016; Zhu, Zhuang, Wong, Cummins, & Tedeschi, 2017). Given the popularity of e-cigarettes, it is important to know which individuals use e-cigarettes, why, and whether different reasons for use are associated with continued use of e-cigarettes and smoking cessation.

As discussed in Chapter 1, one study in the US showed that vapers are more likely to be male and have a lower income (Levy et al., 2017). Additionally, smokers and ex-smokers who had quit in the last 3 years were more likely to have used e-cigarettes than those who had not, but a lower percentage of current and regular vapers were current smokers compared to ever vapers (Levy et al., 2017). In the UK, current and non-current vapers differ in terms of socio-economic status, number of cigarettes smoked per day and quit attempts in the last year (Brown et al., 2014). However, there is limited research exploring whether these differences are observable among young adults in the UK.

Among adults (18+ years) in Great Britain, the primary reasons for vaping are related to smoking cessation (Action on Smoking and Health, 2019a). Currently, there is limited evidence of why young adults in Great Britain vape, but evidence from elsewhere in the world suggests that their reasons for vaping may differ from older adults. Evidence from the US suggests that adolescents and young adults vape primarily out of curiosity (Kong et al., 2015) or because their friends/family vape (Tsai et al., 2018). In a study of South Korean adolescents who had ever smoked and ever vaped, the most common reason for vaping among infrequent users was out of curiosity, but the most common reasons given for frequent vaping were to quit smoking and being able to vape indoors (Lee et al., 2017).

Vaping to quit smoking was also associated with continued vaping among US middle and high school students (Bold et al., 2016). Although some studies have explored associations between reasons for vaping and continuation or discontinuation of vaping and smoking (Bold et al., 2016; Nicksic et al., 2019; Saddleson et al., 2016; Yong et al., 2019), there is limited evidence from the UK. Furthermore, young adults in their mid-twenties are a relatively understudied subgroup in research on reasons for vaping; most research either focusses on school children, students or older adults.

In this study, I aimed to explore the characteristics of young adults who are vaping or have previously vaped, and whether different reasons for vaping are associated with continued vaping and smoking. Specifically, I sought to investigate whether different retrospectively recalled reasons for vaping by 23 years are associated with vaping and/or smoking one year later among a UK cohort of young adults.

3.3 Methods

3.3.1 Study Population

Young adults enrolled in ALSPAC formed the study sample. The profile of the cohort has previously been described in two publications (Boyd et al., 2013; Fraser, Macdonald-Wallis, Tilling, Boyd, Golding, Smith, et al., 2013) and the phases of enrolment are described in more detail in the cohort profile update (Northstone et al., 2019). The total sample size for the cohort is 15,454 pregnancies, resulting in 15,589 foetuses (Figure 3.1). Please note that the study website contains details of all the data that is available through a fully searchable data dictionary and variable search tool (http://www.bristol.ac.uk/alspac/researchers/our-data/). ALSPAC study data from 22 years onwards were collected and managed using REDCap electronic data capture tools hosted at the University of Bristol (Harris et al., 2009).

The questionnaires on vaping and smoking at 23+ years of age were completed by 4,222 study children (from here on in described as young adults) and 3,241 of these also completed the questionnaires at 24+ years of age. The 23+ questionnaire was completed by the young adults between the ages of 23 and 25 and the 24+ questionnaire was completed a year later when the young adults were between 24 and 26 years. For ease, I will refer to these time points as 23 years and 24 years respectively. Figure 3.1 displays the recruitment process in detail.

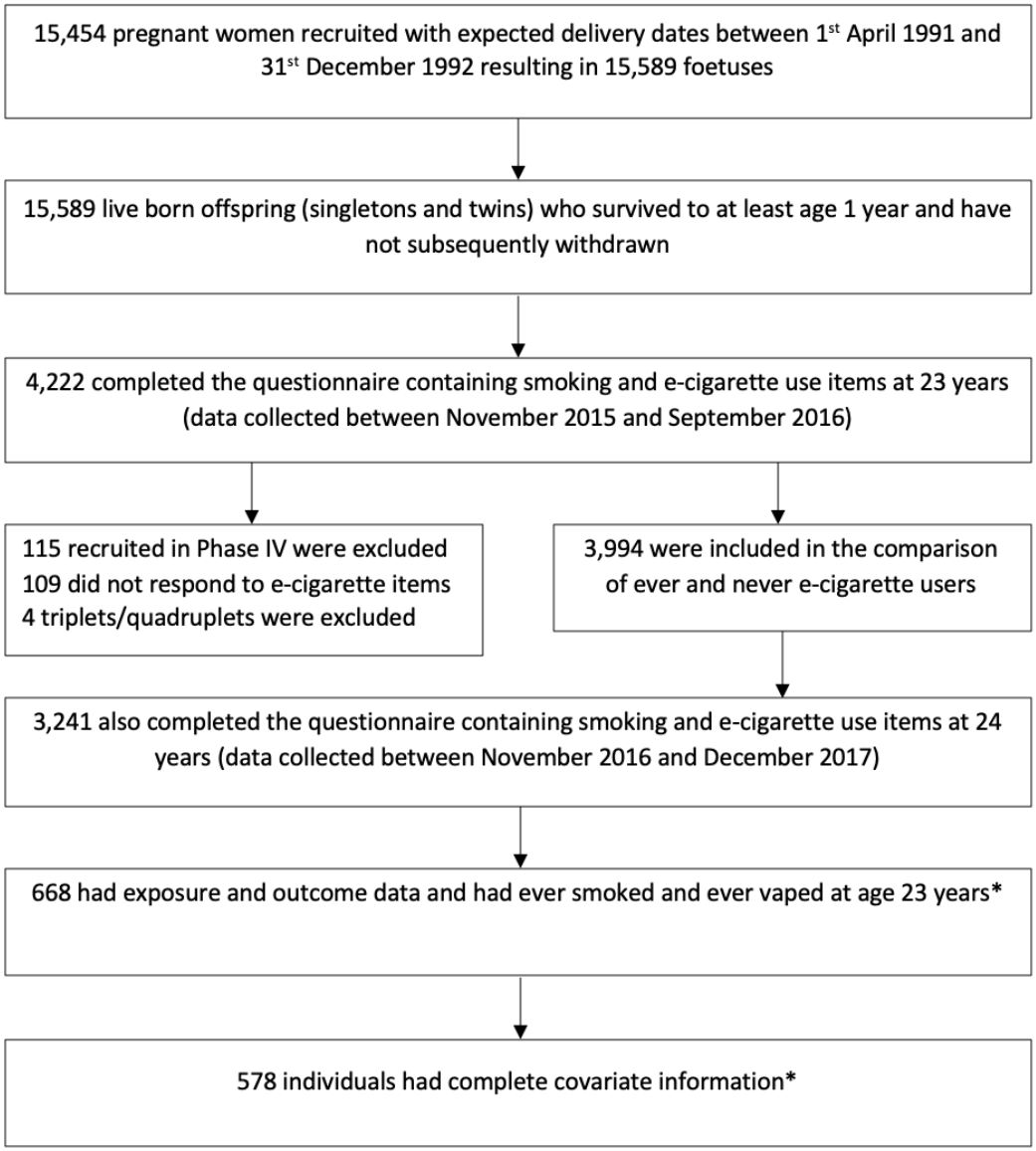


Figure 3.1. Flow chart depicting the process of data inclusion in the analysis of the associations between reasons for vaping at 23 years and vaping and smoking at 24 years.

*Figures shown for ever vapers and ever smokers, for ever vapers and prior smokers n = 412 and n = 360 respectively.

Ethics approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Informed consent for the use of data collected via questionnaires and clinics was obtained from participants following the recommendations of the ALSPAC Ethics and Law Committee at the time.

3.3.2 Measures

3.3.2.1 Participant characteristics. A range of behavioural and lifestyle factors were measured in order to compare the characteristics of: (i) ever vapers (who have vaped at least once) versus never vapers; and (ii) former vapers (who have vaped at least once but reported not currently vaping at 23 years) versus self-reported current vapers. Where these characteristics were measured at multiple timepoints in ALSPAC, I selected the variable which measured the characteristic nearest to the exposure measure at 23 years. Maternal smoking in pregnancy was recorded at 18 weeks gestation and was based on whether the mother responded yes to smoking in the first 3 months of pregnancy. Body mass index (BMI) was measured in clinic at 17 years and was calculated from self-reports of height and weight to the nearest one decimal place using the calculation: weight(kg)/height(m)^2. Measures of risk-taking were taken at 20 years (alcohol, tobacco, cannabis and other drugs and gambling). Alcohol use was measured using the Alcohol Use Disorders Identification Test (Babor, Higgins-Biddle, Saunders, & Monteiro, 2001). Tobacco use was self-reported in response to a question asking whether they had ever smoked a whole cigarette (response options: yes/no). Cannabis use was measured via a question asking if the young person had tried cannabis (response options: yes/no). Other drug use was determined by a response of one or more when asked how many other illicit drugs the young person had ever used. Education/employment status and parenthood status (i.e., whether the young person had become a parent or not) were measured at 21 years. Education and employment status were gathered via a question asking the young person if they were currently in education, training or employment (response options: yes/no). The young adults were asked whether they were a parent at 21 years old. Responses of 'Yes – biological' and 'Yes – non-biological' were recoded to 'yes' to create a binary variable (yes/no). Mental health factors, measured at 21 years, were anxiety and low mood in the past 4 weeks. Anxiety was measured using the General Anxiety Disorder (GAD-7) questionnaire (Spitzer, Kroenke, Williams, & Lowe, 2006). Low mood was measured as the amount of

time in the 4 weeks prior to completing the questionnaire at 21 years that the young person had felt downhearted and depressed. Response options included: 'All of the time', 'Most of the time', 'Some of the time', 'A little of the time', and 'None of the time'. This was recoded into a binary variable of 'None of the time' versus all other responses. Factors relating to vaping (age first vaped, frequency of use, duration of use, device type used) were measured at 23 years (full questions and answer options provided are presented in Appendix 4). Vaping and smoking status were measured at both 23 and 24 years (questions and answer options are described in 3.3.2.2 and 3.3.2.3).

3.3.2.2 Exposure. Exposure was measured at 23 years. To determine whether questionnaire respondents had ever vaped, they were asked "Have you ever used/vaped an electronic cigarette (e-cigarette) or other vaping device?". To determine whether they had ever smoked, they were asked "Have you ever smoked a whole cigarette (including roll-ups)?". To determine whether respondents were regular smokers just prior to vaping, they were asked "Did you smoke tobacco regularly just before you started using electronic cigarettes/vaping devices?". In a multiple-choice question at 23 years, respondents who had ever vaped before were asked "what are/were your reasons for using electronic cigarettes/vaping devices?" and instructed to cross all answers that applied. Seven response options were available: "To help me quit smoking", "To help me cut down on the number of cigarettes I smoke", "To help me with cravings in situations where I cannot smoke e.g. travel, indoors", "Pleasure", "Curiosity", "Friends use them" and "Other". The "Other" option was rarely selected and thus was not included in this analysis.

3.3.2.3 Outcome. At 24 years, e-cigarette outcome data was collected via questionnaire on current vaping. Current vaping was self-reported by the respondent in response to the question "Do you currently use/vape e-cigarettes or other vaping devices?". To determine smoking status at 24 years, the young adults were asked "Have you smoked any cigarettes in the past 30 days?".

3.3.2.4 Potential confounders. Various demographic factors were measured which have previously been shown to impact the likelihood of vaping and smoking and could potentially influence the reason given for vaping (Hartwell et al., 2017; Hedman et al., 2018; Hiscock, Bauld, Amos, Fidler, & Munafo, 2012; Jamal et al., 2016; Karlsen,

Millward, & Sandford, 2012). Sex was recorded at birth. Parental socioeconomic position (SEP) was recorded at 18 weeks gestation and was assessed via parental occupational status (based on the higher of the mother or partner's occupational social class using the 1991 British Office of Population and Census Statistics classification). Ethnicity was recorded at 32 weeks gestation and was classified as white or non-white. The young person's ethnic background was defined as non-white if their mother responded that she or her partner was any other race or ethnic group than white.

3.3.3 Statistical Analysis

3.3.3.1 Differences in participant characteristics. Differences between never and ever vapers as well as former and current vapers were assessed using a χ^2 test (binary outcomes) or t-test (continuous outcomes).

3.3.3.2 Reasons for e-cigarette use and smoking/vaping behaviour. In a series of logistic regressions, I explored the association between retrospectively reported reasons for vaping by 23 years and vaping and smoking continuation at 24 years. Each of the six reasons for vaping were analysed individually as a binary variable (indicated/not indicated as a reason for use). Vaping and smoking status were treated as binary variables (i.e., current vaper versus not current vaper and smoker versus not current smoker). Firstly, I explored the association between reasons for vaping (retrospectively reported at 23 years) and current vaping at 24 among ever vapers at 23 years. I planned to stratify this analysis by ever smokers and never smokers at 23 years. Secondly, I explored the association between reasons for vaping (retrospectively reported at 23 years) and current smoking at 24 among ever vapers at 23 who were regular smokers just prior to first e-cigarette use. These regressions were analysed with and without adjustment for demographic factors (sex, ethnicity, parental SEP, and age in months at 23-year questionnaire).

3.3.3.3 Additional analysis. As it is not clear whether individuals were dual using products in the main analysis (i.e., when separately observing continued vaping and continued smoking), I further explored the association between reasons for vaping and later vaping and smoking status using multinomial logistic regression. Vaping and smoking status were categorised into four groups: current smoker (smoking but not vaping at 24 years), dual user (both vaping and smoking at 24 years), current vaper

(vaping but not smoking at 24 years), or neither user (neither vaping nor smoking at 24 years). Analyses were adjusted for demographic factors (sex, ethnicity, parental SEP, and age in months at 23-year questionnaire). Analyses were restricted to 1) ever vapers who had ever smoked at 23 years and 2) ever vapers at 23 years who had smoked regularly just prior to vaping. I also planned to restrict the analyses to ever vapers who had never smoked at 23 years.

3.3.3.4 Multiple imputation. A majority of the young adults who completed the questionnaire on vaping at 23 years and stated they had ever vaped and ever smoked also had complete outcome data (71%) and complete covariate data (62%). I used multiple imputation to increase the sample size available for analysis and minimise bias due to attrition. There were no clear differences between the young adults who did and did not complete the 24-year questionnaire in terms of reasons for use, other than those who vaped out of curiosity were less likely to have missing outcome data at 24 years ($p = 0.003$). Current vapers ($p < 0.001$) and current smokers ($p < 0.001$) at 23 years were more likely to have missing outcome data at 24 years which indicates that it is plausible that the data are missing not at random (Jakobsen, Gluud, Wetterslev, & Winkel, 2017). As multiple imputation is not advised where data are missing not at random (Jakobsen et al., 2017), I have been conservative in my multiple imputation and only imputed missing covariate data. All those included in the analysis had complete exposure and outcome data. Adjusted analyses were repeated using multiply imputed data. Multiple imputation is a recommended method to account for missing data (Sterne et al., 2009). The multiple imputation by chained equations procedure was completed using the ICE package in Stata 15.1 which created 100 datasets with 20 cycles. I imputed missing covariate data for the 668 young adults who completed the questionnaires at 23 and 24 years and responded that they had ever smoked and ever used an e-cigarette. Further details of the imputation, including the auxiliary variables used, can be found in Appendix 5.

3.3.3.5 Minimum detectable effect. A power calculation indicated that I had sufficient power (90%) to detect a minimum odds ratio of 1.51. The calculation assumed a two-tailed hypothesis ($\alpha = 0.05$) and was based on a logistic regression of vaping to quit smoking by 23 years and vaping status (vaping versus non-vaping) at 24 years**.** In the control group of 440 young adults who had not cited ‘to quit smoking’ as a reason for

vaping, 45 were vapers and 395 were non-vapers at 24 years. Therefore, the calculation was based on an 10% probability of an event in the control group.

3.4 Results

3.4.1 Characteristics of Vapers at Age 23

The characteristics of the young adults are described in Table 3.1, grouped into never vapers (n = 3,013), and ever vapers (n = 981); ever vapers were also grouped into former vapers (n = 814) and current vapers (n = 167). A higher percentage of ever vapers were male, of lower parental SEP at birth, and had a mother who smoked in pregnancy. A higher percentage of ever vapers engaged in other potentially addictive or harmful behaviours; more ever vapers were harmful or hazardous drinkers and had ever used drugs or gambled than never vapers. Ever vapers were also more likely to report anxiety or low mood, smoking by the age of 20 or 23 years, and be current, weekly or daily smokers than never vapers at 23 years. Vaping among never smokers was rare; only 5% of ALSPAC participants who had ever vaped had never smoked, and less than 1% of ever vapers defined themselves as a current vaper who had never smoked at 23 years. There were no clear differences between these groups in terms of ethnicity, BMI, unemployment or parenthood. On average, participants were 23 years old across all groups at the initial questionnaire. The age at which ever vapers first vaped was similar among both current and former vapers (median = 22 years of age). There were few clear differences between former and current vapers. Current vapers were more likely to have had lower parental SEP at birth, report anxiety and have smoked by the age of 23 years but were less likely to be hazardous or harmful alcohol users than former vapers.

Table 3.1. Characteristics of the study population for never, former and current e-cigarette users at 23 years (N = 3,994).

				Ever vapers		
Characteristic	Never vapers (n = 3,013) N (%*)	Ever vapers (n = 981) N (%*)	*p*-value	Former (n = 814) N (%*)	Current (n = 167) N (%*)	*p*-value
Female	2005 (67%)	599 (61%)	.002	502 (62%)	97 (58%)	.387
Parental SEP (manual)	1005 (38%)	372 (43%)	.007	297 (42%)	75 (50%)	.067
Ethnicity (non-white)	116 (4%)	33 (4%)	.509	26 (4%)	7 (5%)	.512
Mother smoked in pregnancy	380 (13%)	205 (22%)	<.001	165 (22%)	40 (25%)	.314
Harmful/hazardous alcohol use	1033 (49%)	401 (67%)	<.001	348 (69%)	53 (55%)	.008
Cannabis use (ever)	923 (41%)	505 (80%)	<.001	424 (80%)	81 (79%)	.673
Other drug use (ever)	401 (18%)	299 (50%)	<.001	255 (51%)	44 (46%)	.354
Gambled (ever)	322 (15%)	143 (24%)	<.001	117 (23%)	26 (26%)	.563
Anxiety	730 (36%)	280 (52%)	<.001	229 (51%)	51 (62%)	.052
Low mood	1269 (63%)	381 (73%)	<.001	321 (72%)	60 (74%)	.742
Overweight/obese BMI	420 (17%)	150 (19%)	.087	122 (19%)	28 (23%)	.268
Currently unemployed	151 (7%)	44 (8%)	.428	37 (8%)	7 (9%)	.920
Parenthood status (had a child)	91 (4%)	32 (6%)	.142	24 (5%)	8 (10%)	.108
Ever smoked by 23 years	1467 (49%)	932 (95%)	<.001	767 (94%)	>162 (>98%)	.005
Current smoker at 23 years	355 (12%)	617 (63%)	<.001	510 (63%)	107 (64%)	.716
Weekly smoker at 23 years**	58 (28%)	111 (39%)	.012	89 (37%)	22 (44%)	.383
Daily smoker at 23 years**	150 (42%)	335 (54%)	<.001	277 (54%)	58 (55%)	.940
	Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)	
Age (initial questionnaire)	23.9 (0.5)	23.9 (0.5)	.854	23.9 (0.5)	23.8 (0.5)	.386

*Due to missing data, the percentage of users refers to the number of participant/participants who responded. **Only current smokers were asked this question. SEP = Socioeconomic position. The highest socioeconomic position of the mother or father was coded as manual vs non-manual occupation. Harmful/hazardous alcohol use was defined as a score of 8 or more on the Alcohol Use Disorders Identification Test (AUDIT). Anxiety was defined as scores of 5 or more on the GAD-7 which indicated mild to severe anxiety. Low mood was defined as feeling downhearted and depressed in the last 4 weeks and was not an indicator of clinical diagnosis. Unemployment status was defined as not currently being employed or engaging in any form of full or part-time education or training. Less than five participants reported current vaping but never smoking by 23 years; the exact number and percentage have been omitted to protect the anonymity of these

participants. Differences between never and ever vapers as well as never, former and current vapers were assessed using χ^2 for binary outcomes and t-test for continuous outcomes.

3.4.2 Reasons for Vaping

Participants' reasons for vaping by 23 years are shown in Table 3.2. Current vapers were more likely than former vapers to vape for all reasons except 'out of curiosity', the most popular reason given for ever vaping (51%). Most young adults (56%) selected only one reason for vaping, 23% selected two reasons, 11% selected three reasons, 4% selected four reasons, 1% selected five reasons and less than 1% selected six reasons for vaping. 4% did not select any reasons for vaping. Due to small numbers of never smokers who had tried vaping at 23 years (n = 47), I did not conduct any inferential statistics comparing current or former vapers to never smokers who have ever vaped. None of the never smokers vaped for smoking-related reasons by 23 years (as expected), 83% vaped out of curiosity, 17% vaped because their friends did, and 6% vaped for pleasure.

Table 3.2. Reasons for e-cigarette use among former and current vapers by 23 years (N = 981).

Reasons for vaping by 23	Former vapers (n = 814) N (%)	Current vapers (n = 167) N (%)	*p*-value	Ever vapers (n= 981) N (%)
Vaped to quit smoking	222 (27%)	113 (68%)	<0.001	335 (34%)
Vaped to cut down number of cigarettes smoked	161 (20%)	68 (41%)	<0.001	229 (23%)
Vaped to help with cravings in situations when unable to smoke	64 (8%)	8 (25%)	<0.001	72 (7%)
Vaped for pleasure	119 (15%)	54 (32%)	<0.001	173 (18%)
Vaped out of curiosity	469 (58%)	32 (19%)	<0.001	501 (51%)
Vaped because friends used them	186 (23%)	22 (13%)	0.005	208 (21%)

3.4.2.1 Reasons for vaping and smoking/vaping behaviour. Due to small numbers of never smokers who had tried vaping at 23 years (n = 47), analyses were restricted to ever-smokers who had ever vaped and were not stratified by ever smoking status. Almost all of the never smokers who had vaped by 23 years did not use either e-cigarettes or cigarettes at 24 years. The study sample consisted of 668 young adults who had completed both questionnaires and had ever vaped and smoked by 23 years. 412 of these young adults were regular smokers just prior to vaping. The median time between questionnaire completion at 23 years and 24 years was 12 months. The age at which the young adults first vaped ranged from 17 to 24 years (SD = 1; median = 22) and on average occurred 2 years before questionnaire completion. Table 3.3 displays vaping characteristics for vapers at 23 years. At 24 years, 49% of the young adults (n = 330) were current smokers who were not currently vaping. Dual use (n = 62) and current vaping (n = 47) were less common. A substantial proportion of the young adults were neither users at 24 years (n = 229). Of the 412 regular smokers just prior to vaping, 59% were current smokers at 24 (n = 244), and dual use (n = 51), current vaping (n = 35), and neither use (n = 82) were less common.

Table 3.3. Prevalence of e-cigarette use characteristics among young adults who had ever smoked and ever vaped at 23 years (N = 668).

E-cigarette use characteristic	N (%)
Current use at 23	105 (16%)
Current smoker and vaper (dual user)	67 (10%)
Current vaper (not current smoker)	38 (6%)
Currently vapes at least monthly	97 (15%)
Currently vapes and has vaped for a month or longer	91 (14%)
Former use by 23	563 (84%)
Used to vape at least once a month but not a current user	195 (29%)
Used an e-cigarette for a month or longer in the past but not a current user	144 (22%)
Device types used by 23	
Ever used a 1st generation device	200 (30%)
Ever used a 2nd generation device	340 (51%)
Ever used a 3rd generation device	410 (61%)

3.4.2.2 Reasons for vaping and associated smoking/vaping behaviour. The unadjusted and unimputed adjusted results of the logistic regressions are shown in Appendix 6 and Appendix 7. The imputed adjusted results are shown in Table 3.4. The results were consistent, with all associations in the same direction with a similar magnitude.

Vaping to quit smoking by 23 years was associated with higher likelihood of continuing to vape at 24 years (odds ratio [OR] = 3.51, 95% confidence interval [95% CI] = 2.29 to 5.38, $p < 0.001$) and a lower likelihood of continuing to smoke (OR = 0.50, 95% CI = 0.32 to 0.78, $p = 0.002$) at 24 years. Vaping to cut down the number of cigarettes smoked by 23 years was associated with a higher likelihood of continuing to vape at 24 years (OR = 2.90, 95% CI = 1.87 to 4.50, $p < 0.001$) and a higher likelihood of continuing to smoke at 24 years (OR = 1.62, 95% CI = 1.02 to 2.58, $p = 0.041$). Vaping to curb cravings for cigarettes by 23 years (OR = 4.35, 95% CI = 2.57 to 7.37, $p < 0.001$) and for pleasure by 23 years (OR = 3.22, 95% CI = 2.01 to 5.15, $p < 0.001$) were also associated with higher likelihood of continuing to vape at 24 years. Vaping out of curiosity by 23 years was associated with lower likelihood of continuing to vape at 24 years (OR = 0.41, 95% CI = 0.26 to 0.63) and a higher likelihood of continuing to smoke at 24 years (OR = 1.66, 95% CI = 1.04 to 2.65, $p = 0.035$). Vaping because friends vaped by 23 years was associated with higher likelihood of continuing to smoke at 24 years (OR = 1.78, 95% CI = 0.95 to 3.36, $p = 0.073$).

Table 3.4. Associations between reasons for vaping by 23 years and current vaping at 24 years among ever vapers and ever smokers and current smoking at 24 years among ever vapers and prior smokers.

Reason for vaping by 23 years	Vaping at 24 years among ever vapers and ever smokers by 23 years (n=668)				Smoking at 24 years among ever vapers by 23 years and regular smokers just prior to vaping* (n=412)			
	N (%)	OR	95% CI	*p*-value	N (%)	OR	95% CI	*p*-value
To quit smoking	228 (34%)	3.51	2.29, 5.38	<.001	217 (53%)	0.50	0.32, 0.78	.002
To cut down	166 (25%)	2.90	1.87, 4.50	<.001	158 (38%)	1.62	1.02, 2.58	.041
To curb cravings	75 (11%)	4.35	2.57, 7.37	<.001	70 (17%)	0.84	0.47, 1.49	.553
Pleasure	118 (18%)	3.22	2.01, 5.15	<.001	57 (14%)	0.88	0.47, 1.65	.685
Curiosity	351 (53%)	0.41	0.26, 0.63	<.001	158 (38%)	1.66	1.04, 2.65	.035
Friends used them	153 (23%)	0.63	0.37, 1.09	.10	71 (17%)	1.78	0.95, 3.36	.073

The analyses were run on multiply imputed data for individuals who ever smoked and ever vaped. Both analyses adjusted for demographic factors (sex, ethnicity, socioeconomic position, and age in months at 23-year questionnaire). *This analysis was restricted to those who had reported that they smoked regularly just before they started vaping to explore continuation/discontinuation of smoking. *Note:* OR = Odds ratio.

3.4.2.3 Additional analysis. The results of the additional analysis which explored the association between reasons for vaping and smoking and vaping status (current vaper, current smoker, dual user or neither user) are displayed in Table 3.5. It is difficult to interpret the results for smoking-related reasons for vaping among the full sample given that non-smokers could not provide a reason relating to quitting or cutting down cigarettes. Thus, associations between smoking-related reasons and smoking and vaping status are only meaningful in the restricted sample of smokers just prior to vaping. The analyses supported the main analyses and also showed that vaping for smoking-related reasons and pleasure were associated with higher likelihood of dual use compared to current smoking at 24 years. Additionally, vaping to quit smoking (among those who were smoking just prior to vaping) and vaping for pleasure (among the full sample) were associated with increased likelihood of neither use compared with current smoking at 24 years. Among those who were smokers just prior to first e-cigarette use, vaping because friends vaped was associated with lower likelihood of being a neither user compared to a current smoker at 24 years. Vaping out of curiosity was associated with lower likelihood of being a dual user and higher likelihood of being a neither user compared to being a current smoker at 24 years among the full sample. However, for those who were smokers just prior to first e-cigarette use, vaping out of curiosity was only associated with a lower likelihood of being a current vaper compared to a current smoker at 24 years.

Table 3.5. Associations between reasons for vaping by 23 years and vaping and smoking status at 24 years among ever vapers and ever smokers as well as ever vapers and smokers just prior to starting to vape.

Reason for e-cigarette use		Full sample (n=668)				Smokers just prior to vaping (n=412)			
Smoking/vaping behaviour		(Yes/No)	aRRR	95% CI	*p*-value	(Yes/No)	aRRR	95% CI	*p*-value
To quit		228/440				217/195			
	Current smoker	114/216	1	(ref)	(ref)	108/136	1	(ref)	(ref)
	Current dual user	36/26	2.74	1.57, 4.78	<.001	33/18	2.36	1.25, 4.44	.008
	Current vaper	28/19	2.82	1.50, 5.31	.001	27/8	4.35	1.88, 10.03	.001
	Neither user	50/179	0.53	0.36, 0.78	.001	49/33	1.83	1.09, 3.06	.021
To cut down		166/502				158/254			
	Current smoker	95/235	1	(ref)	(ref)	91/153	1	(ref)	(ref)
	Current dual user	35/27	3.35	1.90, 5.92	<.001	32/19	2.98	1.57, 5.66	.001
	Current vaper	12/35	0.87	0.43, 1.76	.69	11/24	0.81	0.38, 1.76	.60
	Neither user	24/205	0.29	0.18, 0.47	<.001	24/58	0.71	0.41, 1.24	.23
To curb cravings		75/593				70/342			
	Current smoker	30/300	1	(ref)	(ref)	28/216	1	(ref)	(ref)
	Current dual user	44/18	5.66	2.94, 10.90	<.001	21/216	5.58	2.77, 11.25	<.001
	Current vaper	8/39	2.00	0.84, 4.73	.12	7/28	2.07	0.81, 5.28	.13
	Neither user	14/215	0.70	0.37, 1.35	.29	14/68	1.73	0.85, 3.51	.13
Pleasure		118/550				57/355			
	Current smoker	40/290	1	(ref)	(ref)	23/221	1	(ref)	(ref)
	Current dual user	23/39	4.27	2.29, 7.98	<.001	17/34	4.83	2.28, 10.22	<.001
	Current vaper	15/32	3.35	1.64, 6.85	.001	6/29	2.09	0.76, 5.74	.15
	Neither user	40/189	1.51	0.93, 2.44	.094	11/71	1.55	0.71, 3.41	.27
Curiosity		351/317				158/254			
	Current smoker	163/167	1	(ref)	(ref)	101/143	1	(ref)	(ref)
	Current dual user	28/34	0.78	0.45, 1.37	.39	22/29	1.01	0.54, 1.87	.99

	Current vaper	10/37	0.28	0.13, 0.59	<.001	5/30	0.24	0.09, 0.64	.005
	Neither user	150/79	1.94	1.36, 2.77	<.001	30/52	0.81	0.48, 1.38	.44
Friends used them		153/515				71/341			
	Current smoker	78/252	1	(ref)	(ref)	50/194	1	(ref)	(ref)
	Current dual user	10/52	0.61	0.30, 1.27	.19	7/44	0.60	0.25, 1.42	.24
	Current vaper	8/39	0.70	0.31, 1.57	.39	5/30	0.64	0.24, 1.75	.39
	Neither user	57/172	1.06	0.72, 1.58	.75	9/73	0.47	0.22, 1.01	.052

The models were run on multiply imputed data for individuals who ever smoked and ever vaped (Model 3) and individuals who had been regularly smoking prior to vaping (Model 6). Both models adjusted for demographic factors (sex, ethnicity, socioeconomic position, and age in months at 23-year questionnaire). *Note:* aRRR = adjusted relative risk ratio. Ref = reference category.

3.4.2.3 Exploratory analysis. To further investigate the association between vaping out of curiosity and later neither use (among the full sample only), I ran some exploratory analyses. Specifically, I was interested in whether the association may reflect a group of 'experimenters' who try smoking and vaping but do not continue to smoke or vape. I explored the association between vaping out of curiosity by 23 years and current vaping (versus non-current vaping) and current smoking (versus not current smoking) at 24 years among those who were not regular smokers just prior to vaping (n = 251). Vaping out of curiosity was associated with a lower likelihood of currently vaping at 24 years among young adults who had not regularly smoked just prior to vaping (aOR = 0.24, 95% CI 0.10 to 0.61, p = 0.002) but was not clearly associated with current smoking at 24 years (aOR = 0.63, 95% CI 0.34 to 1.16, p = 0.144).

3.5 Discussion

The results indicate that there are substantial differences in characteristics between young adults who have never and ever vaped at 23. Vaping out of curiosity and to quit smoking were common reasons for vaping among young adults in this UK-based sample. Five out of the six reasons for vaping by 23 years included in this study were strongly associated with continued vaping at 24 years; vaping for reasons related to smoking (to quit, cut down, and curb cravings) and vaping for pleasure were associated with a higher likelihood of continued vaping whereas vaping out of curiosity was associated with a lower likelihood of continued vaping. Four out of six reasons for vaping by 23 years were associated with continued smoking at 24 years; vaping to quit smoking was associated with a lower likelihood of continued smoking whereas vaping to cut down, out of curiosity and because friends vaped were associated with higher likelihood of continued smoking.

Similar to previous findings (Brown et al., 2014), I found that vapers were more likely to have a lower parental SEP at birth than never vapers. Vapers were also more likely to report risk-taking behaviours such as drug use and gambling and were more likely to report mental health issues. Smoking has previously been associated with similar characteristics (Hiscock et al., 2012; Jamal et al., 2016; Lai, Lai, Page, & McCoy, 2000; Minichino et al., 2013). This similarity could indicate a common liability for both behaviours (i.e., the same factors increase the likelihood of engaging in both behaviours)

which could explain why not all young adults who vape go on to quit smoking. Alternatively, smoking could mediate the relationship between risk-taking and vaping; most vapers in this cohort smoked prior to vaping and some started vaping to help them quit smoking. In line with previous findings (Levy et al., 2017), vapers were more likely to be ever, weekly or daily smokers at 23 years, and vaping among never smokers was rare. Even though 52% of the sample did not provide a smoking-related reason for vaping and I cannot determine order of product use for all participants, it is unlikely that vaping led to smoking in this cohort; only 1% of the young adults were not smokers just prior to first using an e-cigarette but stated they had started smoking regularly since using an e-cigarette. However, this figure may include ex-smokers who took up vaping to avoid relapsing to smoking (and therefore were not regularly smoking *just* before starting to use an e-cigarette).

My results imply that vaping to quit smoking may facilitate young adult's quit attempts; young adults who smoked just prior to vaping who vaped for this reason were less likely to continue to smoke at 24 years and were more likely to be neither users than current smokers compared with those who did not vape for this reason. This is in line with the growing amount of evidence that vaping can facilitate smoking cessation (Hartmann-Boyce et al., 2020). My findings also support previous findings among middle and high school students (Bold et al., 2016) and college students in the US (Saddleson et al., 2016) which found that vaping to quit smoking was associated with higher likelihood of continued vaping compared with not vaping for this reason. For some individuals, vaping to quit smoking could encourage the continued use of e-cigarettes rather than quitting nicotine products entirely. If these people aim to be completely nicotine free, they may need additional help to quit vaping or reduce the levels of nicotine in their e-liquid once they have successfully quit smoking.

Although vaping to cut down smoking is similar in terms of motive, the associated behaviour was quite different from vaping to quit smoking. Those who were regular smokers just prior to starting to vape and vaped to cut down were more likely to still be smokers at 24 years than those who did not vape to cut down. Similar to previous findings (Bold et al., 2016) they were also more likely to continue vaping (i.e., be dual users) which suggests that these individuals were not encouraged to quit smoking by vaping to cut down. There are several potential explanations for this difference. First,

these users may not actually intend to quit smoking, and intention may be necessary for e-cigarettes to act as an effective smoking cessation tool. Second, the comparison group includes those who vaped with the sole intention to quit smoking which could potentially mask a decreased likelihood of continuing smoking among those who vaped to cut down. However, adjusting for use to quit smoking did not substantially change the results (OR = 1.62, 95% CI = 1.01 to 1.60). Third, recall bias may be an issue; those who were unable to quit smoking may report that they only intended to cut down. Fourth, as the study was over a relatively short time period, these users may still be in the process of quitting and may eventually quit entirely.

Vaping to curb cravings for cigarettes and vaping for pleasure were less common. There is little evidence to suggest that vaping to curb cravings or for pleasure hampers smoking cessation. Equally, there is little evidence to suggest that vaping for these reasons encourages smoking cessation.

Consistent with previous research conducted in the US, I found that curiosity is the most common reason for vaping among young adults (Kong et al., 2015). In line with previous evidence (Bold et al., 2016), young adults were less likely to continue to vape at 24 years and were more likely to be neither users than current smokers if they had vaped out of curiosity rather than if they had not. As the full sample are *ever* smokers who have *ever* vaped, this may highlight a group of 'experimenters' who try both vaping and smoking but do not continue either behaviour. This is supported by my exploratory analysis; there was no clear evidence of an association with vaping out of curiosity and later smoking among those who were not regular smokers just prior to initially vaping. Similarly, for those who had smoked less than 100 cigarettes in their lifetime, curiosity was a common reason for use in a study of young adults in the US (Biener, Song, Sutfin, Spangler, & Wolfson, 2015). This is also consistent with what I hypothesised in the discussion of Chapter 2; although there is a strong association between ever vaping among never smokers and subsequent smoking, this could be due to a shared liability (e.g., propensity to risk-taking) which leads young adults to experiment with both products but not lead to continued smoking. I further test this hypothesis in Chapter 4. Interestingly, those who had been regular smokers just prior to starting to vape and who did so out of curiosity were more likely to still be smoking at 24 years and were less likely to be current vapers than be current smokers compared to those who did not vape

out of curiosity. This implies that smokers who vape out of curiosity do not want to quit or are less able to quit smoking compared to those who vape for other reasons (e.g., to quit smoking). Ultimately, the evidence suggests that vaping out of curiosity is not associated with transitions in smoking/vaping status; vaping out of curiosity was negatively associated with continued vaping, and there was only an association with later smoking among those who were already regularly smoking just prior to vaping.

These results provide strong evidence for associations between reasons for e-cigarette use and later vaping and smoking status using a large UK-based young adult cohort while adjusting for a range of potential confounders. However, further research is needed to fully explore the potential implications of these findings using methods which can support stronger causal inference. For example, to explore whether intention to quit smoking is necessary for young adults to quit smoking using e-cigarettes, a randomised controlled trial could be designed whereby young adults receive instructions to either try to cut down the number of cigarettes smoked or quit all together. Comparing smoking cessation rates between the two conditions would allow us to infer with greater confidence whether intention to quit is necessary for effective smoking cessation among young adults using e-cigarettes. However, instructing an individual to quit may not be sufficient for them to actually intend to quit, so it may be a poor proxy for intention. Therefore, the results may not accurately reflect intention to quit but reflect instruction to quit (i.e., an intention to treat model). Nevertheless, if the findings of such a study were to corroborate the current findings, it would provide stronger support for interventions aiming to encourage smokers to quit smoking with the assistance of e-cigarettes.

Despite utilising a rich data source (ALSPAC), the study was limited by the measures included. This was particularly an issue for the exposure; the reasons for vaping were self-reported retrospectively. This measure could suffer from recall bias as individuals may have started vaping years prior to completing the questionnaire. For example, young adults may be less likely to state that they vaped to quit smoking if they were unsuccessful in their quit attempt. Also, the reasons for vaping chosen were not exclusive for each young person due to the multiple-choice format of this question. As with many previous studies, a multiple-choice item was used to determine reasons for vaping where open questions and qualitative analysis may have uncovered other

potential reasons for vaping which are omitted from this analysis. The young adults were also provided with an 'other' option in the questionnaire where they could provide open answer responses, however, this was rarely selected (4% of the sample) or selected without another reason being specified (2%) so was not included in the analysis. Vaping for flavour-related reasons is common in the US (Landry et al., 2019) but 'flavours' was not an option in the 23-year questionnaire. 'Flavours' was included in the 24-year questionnaire (selected by 17% of respondents). In cross-sectional analysis, there was evidence that vaping for flavours by 24 years was associated with increased likelihood of current vaping (aOR = 1.70, 95% CI 1.15 to 2.53), but was not clearly associated with current smoking (aOR = 1.06, 95% CI 0.59 to 1.95). Additionally, the answer options that were provided may have primed a response in the young adults which was biased towards smoking cessation as three of the six were related to smoking.

Although I employed multiple imputation methods to minimise bias (Sterne et al., 2009) and the ALSPAC response rate has remained consistent in the last eight yearly questionnaires (~4000 respondents), another limitation of this study is that missing data could have introduced selection bias. Evidence suggests that smokers may be less likely to participate in ALSPAC (Taylor et al., 2018); therefore, there may be fewer individuals vaping or vaping for smoking-related reasons than in the full sample.

The timing of this cohort study may impact the generalisability of the findings. E-cigarettes are still a relatively new product compared to cigarettes; consequently, cigarettes were available to this cohort of young adults for a considerable period of their adolescence before e-cigarettes became widely available in 2007. In 2007, the study sample were roughly 17 years old and cigarette initiation has been shown to peak at around 15-16 years of age (Marcon et al., 2018), so it is likely that many of these young adults experimented with cigarettes prior to being exposed to e-cigarettes. Adolescents today are being exposed to both e-cigarettes and cigarettes during adolescence, a key period for experimentation and risk-taking behaviour, and this may have an impact on their reasons for vaping as well as their current vaping and smoking status. Although it would be interesting to observe the association between reasons for vaping and later vaping and smoking status among those who were never smokers when they first vaped, I was unable to accurately identify these individuals with the current available data. It

would also be interesting to observe the associations among never smokers at 23 years but there are too few individuals who had vaped but never smoked at 23 years (who also responded to the questionnaire at 24 years) to conduct any meaningful analysis. Of note, none of these excluded individuals used an e-cigarette for smoking-related reasons by 23 years (as expected), the majority vaped out of curiosity, and almost all of them were neither users at 24 years.

3.6 Chapter Summary

In conclusion, vapers were more likely than never vapers to be male, of lower parental SEP at birth, have poor mental health and engage in other risky behaviours as well as be current weekly or daily smokers. These factors are also associated with smoking which, as suggested in Chapter 2, could indicate a shared liability to both behaviours, or could be due to smoking leading to e-cigarette use. I also found that vaping to quit smoking is associated with increased likelihood of later vaping but decreased likelihood of later smoking among those who were smokers just prior to starting to use e-cigarettes. In contrast, vaping to cut down was associated with continued smoking. This implies that intention to quit smoking may be necessary for young adults to effectively stop smoking using e-cigarettes. Vaping out of curiosity does not appear to lead to continued vaping or smoking among those who were not regularly smoking just prior to vaping. In Chapter 2, I found that e-cigarette use is associated with later smoking and my findings in this Chapter could be consistent with either of my suggested explanations for this association. Curious vapers my not vape enough to become dependent on nicotine, and therefore do not seek other nicotine sources, or smoking and vaping may share a common liability (i.e., a propensity to experiment and engage in risky behaviour). I have further explored these potential shared factors (SEP and risk-taking) while investigating whether there is a shared genetic liability to both smoking and vaping in Chapter 4.

Chapter 4 Does genetic liability to smoking initiation influence e-cigarette use?

This chapter closely resembles sections from the following pre-print:

Khouja, J. N., Wootton, R. E., Taylor, A. E., Davey Smith, G., & Munafò, M. R. (in press). Association of genetic liability to smoking initiation with e-cigarette use in young adults. *PLOS Medicine*.

Khouja, J. N., Wootton, R. E., Taylor, A. E., Davey Smith, G., & Munafò, M. R. (2020). Association of genetic liability to smoking initiation with e-cigarette use in young adults. *MedRxiv*; doi: https://doi.org/10.1101/2020.06.10.20127464.

I took the lead in designing, implementing and interpreting this study with the support of my co-authors. Specifically, MM, RW, AT and GDS provided feedback on the design of the study and during the preparation of the manuscript. I was responsible for the data analysis and was responsible for writing the first draft of the manuscript and editing the manuscript in response to comments from co-authors and reviewers.

4.1 Chapter Overview

In Chapter 1, I described the strong association between e-cigarette use and smoking which could be due to a causal effect of e-cigarette use on smoking, a causal effect of smoking on e-cigarette use, or a common cause (or confounding) between the two behaviours. As seen in Chapter 2, never smokers who use e-cigarettes are more likely to subsequently smoke than those who do not; this could imply that e-cigarette use causes adolescents and young adults to smoke. In contrast, as shown in Chapter 3, a common reason given for using e-cigarettes is to quit smoking; therefore, smoking leads young adults to use e-cigarettes to alleviate symptoms of withdrawal when quitting. Another explanation for the association between e-cigarette use and smoking is that the behaviours share a 'common liability' such as a genetic predisposition to both smoke and vape. In Chapter 3, I found that a higher percentage of ever vapers engaged in other potentially addictive or harmful behaviour than never vapers, suggesting the shared liability may relate to risk-taking. As discussed in Chapter 1, there are no well-powered

genome-wide association studies (GWAS) of e-cigarette use to directly compare the genetic variants associated with e-cigarette use and smoking, but there is a well-powered GWAS of smoking initiation. Using summary results from the smoking initiation GWAS, I create smoking initiation polygenic risk scores (PRS) at varying *p*-value thresholds for young adults (23-26 years) of European ancestry in the Avon Longitudinal Study of Parents and Children (ALSPAC). In this Chapter, I explore the association between smoking initiation PRS and ever e-cigarette use at 24 years in ALSPAC to explore shared genetic influences on e-cigarette use.

4.2 Introduction

As discussed in Chapter 3, the use of e-cigarettes for smoking cessation is common among young adults in the UK (Khouja, Taylor, & Munafò, 2020); therefore, it would be logical to assume that smoking causally influences e-cigarette use in this population. However, some studies have shown an association between e-cigarette use and subsequent smoking among non-smokers (Khouja, Suddell, Peters, Taylor, & Munafo, 2020; Soneji, Barrington-Trimis, Wills, Leventhal, et al., 2017), which suggests the possibility that e-cigarette use may also act as a gateway to smoking (sometimes referred to as the gateway hypothesis), particularly among adolescents. In Chapter 2, in which I described the meta-analysed results of these studies, I found a strong and consistent positive association between e-cigarette use (among never smokers) and later smoking for youth aged 30 years or younger, but I concluded that there is currently insufficient evidence that this association is causal (Khouja, Suddell, et al., 2020). I discussed the distinct possibility that the association may be due to vaping and smoking sharing a common liability. In Chapter 3, I found that a higher percentage of ever vapers engaged in other potentially addictive or harmful behaviour than never vapers, suggesting the shared liability may relate to risk-taking. Understanding more about the nature of the association between smoking and e-cigarette use is vital to inform tobacco control policies that aim to prevent youth smoking initiation (e.g., by restricting access to e-cigarettes). Specifically, it is important to understand whether the association found among adolescents and young adults is causal, or due to other factors that influence both smoking and e-cigarette use independently.

As I previously mentioned in Chapter 1, there is some evidence for a shared genetic liability to both smoking and e-cigarette use (Allegrini et al., 2019). This could indicate a causal relationship in that smoking causes vaping (i.e., vertical pleiotropy), or it could be due to genetic variants that separately influence both behaviours (i.e., horizontal pleiotropy) (Davey Smith & Hemani, 2014). One biologically plausible explanation for a genetic link between smoking and e-cigarette use is that they are both influenced by the same genetic variants that influence an individual's response to nicotine or their nicotine metabolism. However, evidence suggests that some of the genetic influence on smoking initiation is mediated by personality traits, such as risk-taking and impulsivity, that influence (among other things) smoking uptake (Heath, Madden, Slutske, & Martin, 1995). These traits can be observed in childhood (prior to smoking or vaping initiation), particularly among those with externalising disorders (Samek & Hicks, 2014). Using data from an early GWAS of smoking (Tobacco and Genetics Consortium, 2010), Allegrini and colleagues (2019) suggested that the genetic link found in their research between smoking and e-cigarette use may reflect these personality traits.

Using genetic variants, I can explore whether genetic predisposition to smoking is associated with e-cigarette use, and which factors or mechanisms may influence the association. As I discussed in Chapter 1, I would ideally explore the genetic overlap between smoking and e-cigarette use by comparing the genetic variants identified in a GWAS of each behaviour, but at present there are no large, well-powered GWAS of e-cigarette use. However, a GWAS of various smoking behaviours which identified 378 single nucleotide polymorphisms (SNPs) associated with smoking initiation is available (Liu et al., 2019). Using these SNPs, smoking initiation PRS can be created and associations between these PRS and a range of outcomes can be examined.

Causality cannot be inferred from such analyses, but as I explained in Chapter 1, negative control outcomes can be used to inform the overall evaluation of whether an association is causal via a hypothesised route (Lipsitch, Tchetgen Tchetgen, & Cohen, 2010). Triangulating evidence from outcomes where a simple biological pathway from smoking to the outcome is implausible (e.g., gambling), or impossible (e.g., externalising behaviour or socioeconomic position [SEP] in childhood, before smoking has occurred) can aid our understanding of potential pathways by which smoking and vaping may share a genetic predisposition. These potential pathways (displayed in Figure 4.1)

include a biological pathway from smoking to e-cigarette use (i.e., vertical pleiotropy), a shared genetic predisposition which influences smoking and vaping independently (i.e., horizontal pleiotropy), or a genetic liability to a broader, risk-taking phenotype which causes both smoking and vaping (i.e., a shared risk factor). Alternatively, triangulation could aid our understanding of whether an association is due to a shared genetic predisposition between parents and offspring. Where parents share their offspring's smoking initiation predisposition and consequently expose their offspring to cigarette smoke in utero or in childhood, an apparent effect of a child's own genetic variants may be a result of their environment due to their parents' genetic variants. If associations are only found between smoking initiation PRS and e-cigarette use, but not negative control outcomes, this would support the vertical pleiotropy interpretation; however, if an association is also found with negative control outcomes, this would suggest that horizontal pleiotropy is occurring or that shared parent-offspring genetic predisposition may be confounding the association.

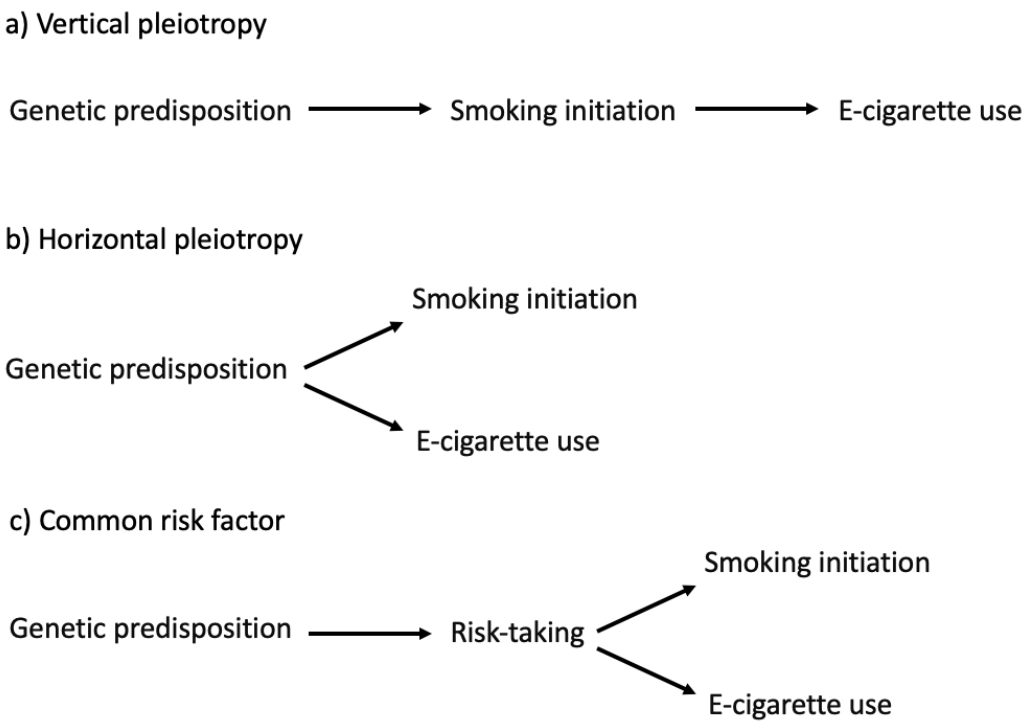


Figure 4.1. Potential models of shared liability for the relationship between genetic predisposition to smoking initiation and e-cigarette use.

Additionally, using varying *p*-value thresholds to create PRS could help to identify the presence of horizontal pleiotropy. Calculating PRS at less strict *p*-value thresholds than the genome-wide significant threshold increases the percentage variance in the phenotype explained by the score, and thus increases power to detect an association. However, using less stringent thresholds will also tend to increase the likelihood of including genetic variants which are related to factors other than the exposure of interest, making the PRS less specific (and may eventually result in PRS which explain less variance in the exposure). The more SNPs included in a PRS, the less likely it is that the effect of each variant on the trait of interest is proportional to the effect of each variant on the exposure, and the more likely it is proportional to the effects on other (horizontally pleiotropic) traits (Hemani, Bowden, & Davey Smith, 2018), increasing the likelihood that any associations found between the PRS and an outcome could be due to horizontal pleiotropy. Triangulating evidence from a variety of thresholds and a variety of outcomes may provide a clearer picture of the true association; associations observed when more stringent PRS thresholds are used could be due to a causal effect of smoking, while associations observed only at less stringent thresholds among negative control outcomes may indicate horizontal pleiotropy.

In this study, I investigated whether smoking initiation PRS are associated with ever use of e-cigarettes in young adulthood. I also aimed to explore any associations with outcomes that are not plausibly biologically related (e.g., gambling) or that precede smoking (e.g., hyperactivity in childhood), to determine whether the association between smoking and e-cigarette use could reflect a broader risk-taking phenotype captured by the smoking initiation PRS.

4.3 Methods

4.3.1 Data Sources

4.3.1.1 GSCAN. The GWAS & Sequencing Consortium of Alcohol and Nicotine use (GSCAN) report summary level statistics from a GWAS of smoking initiation (Liu et al., 2019). This GWAS was based on 1,232,091 participants from 29 cohorts. I obtained

summary statistics, with ALSPAC (N = 11,345) removed, through correspondence with GSCAN to eliminate data overlap with the target sample. Due to data sharing restrictions, 23andMe were also excluded from this data (N = 599,289) leaving a total sample size of 621,457. Smoking initiation was defined as ever being a regular smoker. The exact definition varied across the cohorts included in the GWAS, with 3 different definitions: 1) Have you smoked over 100 cigarettes over the course of your life? 2) Have you ever smoked every day for at least a month? 3) Have you ever smoked regularly?

4.3.1.2 ALSPAC. The target sample consisted of participants from ALSPAC (Boyd et al., 2013; Fraser, Macdonald-Wallis, Tilling, Boyd, Golding, Davey Smith, et al., 2013). This study recruited pregnant women residing in Avon, UK with expected dates of delivery from 1st April 1991 to 31st December 1992. The phases of enrolment are described in detail in the cohort profile paper and its update (Northstone et al., 2019). A total of 15,454 mothers were recruited, resulting in 15,589 foetuses. Of these, 14,901 were alive at 1 year of age. Genetic data was available for 9,085 young adults and, after samples which did not pass quality control were removed, PRS were created for 7,859 unrelated individuals of European ancestry. Of these individuals, 2,905 also had data for our main outcome at 24 years regarding their vaping behaviour. Sample sizes varied by outcome due to restrictions (e.g., restricting to never smokers) and differing timepoints of measurement (i.e., missing data).

4.3.2 Polygenic Risk Scores

I used summary data from GSCAN (excluding ALSPAC and 23andMe, N = 621,457) to select SNPs associated with smoking initiation. GSCAN report beta coefficients of the change in standard deviation in smoking initiation per one unit of change in genetic risk (i.e., per risk allele). I converted the beta coefficients to log odds ratios and gave each participant a score which indicated the average number of risk alleles (0, 1 or 2 effect alleles) they possessed for the selected SNPs. As some SNPs are more strongly associated with the phenotype (smoking initiation), I multiplied the number of risk alleles by the log odds ratios from the summary statistics (with ALSPAC and 23andMe removed) to create a weighted score. Then, I standardised the scores by transforming them to z-scores. Z-scores indicate how many standard deviations a value lies away from the mean (Field, 2013). Using five *p*-value thresholds ($5x10^{-8}$, 0.0005, 0.005, 0.05, 0.5), I

selected five groups of SNPs to be included in five different PRS for each participant. I used PLINK to create PRS at the $p < 5x10^{-8}$ threshold using the SNPs which met the genome-wide significance threshold in the GSCAN GWAS of smoking initiation (Liu et al., 2019). Using PRSice software, I calculated the PRS at all other thresholds (Euesden et al., 2015). The data I acquired from GSCAN had already been pruned for SNPs with a Minor Allele Frequency (MAF) > 0.001 where at least 10% of the maximum sample size had SNP data available in at least three of the consortium studies. In addition, I pruned the SNPs using PLINK software to ensure low linkage disequilibrium ($r^2 < 0.1$).

4.3.3 Outcomes

4.3.3.1 E-cigarette use. At 24 years (between 2016 and 2017), outcome data was collected via questionnaire on whether participants had ever vaped. Ever use was defined as ever having used/vaped an e-cigarette or other vaping device and was determined by response to the question "Have you ever used/vaped an electronic cigarette (e-cigarette) or other vaping device?". Response options were: yes; no.

4.3.3.2 Smoking. I included self-reported ever smoking and smoking initiation as positive control outcomes (i.e., outcomes for which an association with the exposure is expected). Ever smoking by 24 years was defined as having ever smoked a whole cigarette (including roll-ups) and was determined by response to the question "Have you ever smoked a whole cigarette (including roll-ups)?". Response options were: yes; no. Smoking initiation by 24 years was defined as having smoked 100 or more cigarettes in their lifetime which was determined by response to the question "How many cigarettes have you smoked altogether in your lifetime?". Response options were: less than 5; 5-19; 20-49; 50-99; 100 plus. I recoded these responses into 2 categories: initiated smoking (100 plus cigarettes) and did not initiate smoking (all other response options).

4.3.3.3 Negative controls. I included four negative control outcomes at age 23 and 24 in the analysis: high number of sexual partners, having been in trouble with the law, ever gambling, and enjoying taking risks. These variables were selected on the basis of being related to broad risk-taking behaviour, but where a causal pathway from smoking was not considered biologically plausible. To determine the number of sexual partners the young adults had had, they were asked "All together, in your life so far, how many people have you had sexual intercourse with?". Responses were free recall. I

determined the threshold considered to be a high number of sexual partners at 23 years by using the upper quartile for number of lifetime sexual partners in the ALSPAC sample (11 or more self-reported sexual partners at 23 years). Having been in trouble with the law recently (since 23rd birthday) was self-reported at 24 years and was determined by response to the question "Have any of these happened since you were 23 years old and did they affect you? You were in trouble with the law." Response options were: yes, affected me a lot; yes, moderately affected; yes, mildly affected; yes, but didn't affect me at all; no, did not happen. I recoded these responses into two categories: yes (responses beginning with "yes") and no (response "no, did not happen"). I determined whether a young adult had ever gambled by 24 years using responses to the question "Have you ever participated in any of the form of gambling listed?". The list included: Tickets for the National Lottery; Scratchcards; Tickets for any other lottery; The football pools; Bingo cards or tickets; Fruit slot machines; Virtual gaming machines in a bookmaker's to bet on a virtual roulette, keno, bingo etc.; Table games (roulette, dice or cards) in a casino; Online gambling like playing poker, bingo, slot machine style games, or casino games for money; Online betting with a bookmaker on any event or sport; Betting exchange; Betting on horse races in a bookmakers, by phone, or at the track; Betting on dog races in a bookmakers, by phone, or at the track; Betting on any other event or sport at the bookmakers, by phone or at the venue; Spread-betting; Private betting, playing cards or games for money with friends, family or colleagues; Any other form of gambling. If the participant had ever engaged in any of these activities, I considered them to have ever gambled. Enjoying taking risks at 24 years was determined in response to a question asking to "indicate how much you agree or disagree with [this] statement" with the statement being "I quite enjoy taking risks". Response options were: agree strongly; agree somewhat; disagree somewhat; disagree strongly. I recoded these responses into 2 categories: yes (responses beginning with "agree") and no (responses beginning with "disagree").

I included three negative control outcomes at age 7: hyperactivity, conduct disorder (CD) and oppositional defiant disorder (ODD). These externalising disorders are indicators of impulsivity in childhood and later life (Samek & Hicks, 2014) and were selected on the basis that few (if any) children at this age have smoked, ruling out a causal pathway from their own smoking to these outcomes. Hyperactivity and conduct disorder were measured using the Strengths and Difficulties Questionnaire (SDQ) at 7

years and ODD was measured using the Development and Well-Being Assessment (DAWBA) at 7 years. Scores of 0-5 for hyperactivity are considered 'normal', and above 5 is considered borderline/abnormal (i.e., hyperactive/attention deficit hyperactivity disorder [ADHD]) out of a possible score of 10. Scores of 0-2 out of a possible 10 for conduct disorder (determined by teacher/parent) are considered 'normal', and above are considered borderline/abnormal (i.e., CD). I considered any awkward symptoms (out of a possible 18) before the age of 7 to be abnormal (i.e., ODD).

I also included parental SEP at birth in the analysis, which was assessed using the 1991 Office of Population Censuses and Surveys (OPCS) occupation classification (manual vs non-manual; measured during pregnancy at 32 weeks gestation). This outcome was based on highest occupation of both parents at birth (preceding smoking) and was selected on the basis that it could not possibly be caused by a young person's own smoking.

4.3.4 Statistical Analyses

I carried out all of my analysis in Stata 15.1 (StataCorp, 2017) using logistic regression which I adjusted for age, sex and the first 10 principal genetic components of population stratification. Principal components adjustment is often used in genetic research to address issues of population stratification (Zhao, Mitra, Kanetsky, Nathanson, & Rebbeck, 2018) whereby common differences in allele frequencies exist among sub-populations. I assessed the association between smoking initiation PRS and (i) ever e-cigarette use by age 24, (ii) regular e-cigarette use at age 24, (iii) smoking initiation, and (iv) negative control outcomes (risk-taking behaviours, externalising disorders, and SEP).

I repeated the main analysis using a restricted sample of never smokers (who have smoked less than 100 cigarettes in their lifetime) to explore whether any associations remained when the association could not be caused by the young adult's own smoking. I also restricted the analysis to those who had never smoked a whole cigarette in their lifetime for comparison.

4.4 Results

A total of 378 SNPs were identified as genome-wide significant in the GSCAN GWAS of smoking initiation (Liu et al., 2019), 356 of which were available in ALSPAC. I removed nine SNPs at the pruning stage (removing SNPs with $r^2 > 0.1$ within 50 kb), leaving 347 SNPs for inclusion in the most stringent PRS (*p*-value threshold $p < 5x10^{-8}$). The number of SNPs included in each PRS at the less stringent thresholds are shown in Table 4.1. Of note, for the PRS calculated at these less stringent thresholds, I used the significance level reported in the restricted sample (excluding ALSPAC and 23andMe) summary data.

Table 4.1. *p*-value thresholds and number of SNPs included in polygenic risk scores.

p-value thresholds	N SNPs for smoking initiation
$5x10^{-8}$	347
0.0005	2,808
0.005	10,098
0.05	42,494
0.5	169,364

4.4.1 Characteristics of the sample

Table 4.2 shows the characteristics of the sample; 878 (30%) young adults were self-reported ever e-cigarette users by 24 years and 1,695 (64%) were self-reported ever smokers. Of those who had ever vaped, 95% (n = 830) had ever smoked at least one whole cigarette and 71% (n = 616) had smoked 100 or more cigarettes. Less than 1% of the sample had used an e-cigarette prior to smoking. Self-reported smoking and e-cigarette use were associated with lower parental SEP and having externalising disorders in childhood (Table 4.3). Self-reported smoking and e-cigarette use were also associated with increased odds of engaging in risk-taking behaviours (Table 4.3).

Table 4.2. Characteristics of young adults in ALSPAC.

Characteristic	**N (%)**
Ever used an e-cigarette by 24 (used once or more)	878 (30%)
Regularly used an e-cigarette at 24 (used at least once a month)	150 (5%)
Ever smoked by 24 (1 cigarette or more)	1,695 (64%)
Initiated smoking by 24 (100 cigarettes or more)	972 (33%)
Ever used an e-cigarette but not initiated smoking by 24	262 (13%)
High number of sexual partners at 23*	647 (25%)

Been in trouble with the law since 23rd birthday	69 (2%)
Enjoys taking risks at 24	1,618 (55%)
Ever gambled at 24	2,156 (74%)
Hyperactivity at 7	2,219 (42%)
Conduct disorder at 7	1,199 (22%)
Oppositional defiant disorder at 7	1,868 (35%)
Parental SEP (manual)	1,068 (27%)
	Mean (SD)
Age in months at 24-year questionnaire	298 (6)

Note: Sample sizes varied by characteristic due to differing timepoints of measurement (i.e., missing data).* 11 or more sexual partners, determined using the upper quartile for number of lifetime sexual partners in the ALSPAC sample (11 sexual partners).

Table 4.3. Association between self-reported e-cigarette use and smoking and risk-taking behaviours, socioeconomic indicators and externalising disorders in childhood.

Outcome Exposure	**n**	**OR**	**95% CI**	**p**
Ever used e-cigarettes by 24				
Number of sexual partners at 23*	2013	2.40	1.93, 2.97	<0.001
Been in trouble with the law since 23rd birthday	2891	2.43	1.49, 3.95	<0.001
Enjoys taking risks at 24	2896	1.62	1.37, 1.90	<0.001
Ever gambled at 24	2860	2.05	1.68, 2.51	<0.001
Hyperactivity at 7	2340	1.34	1.12, 1.61	0.002
Conduct disorder at 7	2375	1.24	1.00, 1.53	0.054
Oppositional defiant disorder at 7	2380	1.25	1.04, 1.51	0.017
Parental SEP	2630	1.46	1.13, 1.90	0.004
Ever initiated smoking by 24 (>100 cigarettes)				
Number of sexual partners at 23*	2034	3.21	2.60, 3.96	<0.001
Been in trouble with the law since 23rd birthday	2922	2.12	1.30, 3.44	0.002
Enjoys taking risks at 24	2927	1.78	1.52, 2.09	<0.001
Ever gambled at 24	2890	1.67	1.39, 2.02	<0.001
Hyperactivity at 7	2363	1.22	1.02, 1.45	0.032
Conduct disorder at 7	2396	1.17	0.95, 1.45	0.143
Oppositional defiant disorder at 7	2407	1.23	1.02, 1.47	0.026
Parental SEP	2655	1.59	1.24, 2.05	<0.001

*Low (<11) vs. high (11 or more) number of sexual partners, determined using the upper quartile for number of lifetime sexual partners in the ALSPAC sample (11 sexual partners).

4.4.2 PRS for Smoking Initiation and Self-Reported Smoking

I observed positive associations between smoking initiation PRS and ever smoking (having smoked at least 1 cigarette in a lifetime) by the age of 24 years ($p < 5x10^{-8}$ threshold OR [OR_{10-8}] = 1.25, 95% CI 1.16 to 1.35) and smoking initiation (having smoked at least 100 cigarettes in a lifetime) by the age of 24 years (OR_{10-8} = 1.29, 95% CI 1.19 to 1.39). I found strong associations between PRS and self-reported smoking measures at all *p*-value thresholds (Table 4.4).

Table 4.4. Associations between polygenic risk scores for smoking initiation with ever e-cigarette use, ever smoking and smoking initiation.

Outcome *p*-value threshold	**N**	**OR**	**95% CI**	**p**
Ever e-cigarette use by 24	2,894			
5×10^{-8}		1.24	1.14, 1.34	<0.001
0.0005		1.27	1.17, 1.38	<0.001
0.005		1.36	1.26, 1.48	<0.001
0.05		1.39	1.28, 1.51	<0.001
0.5		1.39	1.28, 1.51	<0.001
Regular e-cigarette use at 24 (at least once a month)	2,894			
5×10^{-8}		1.18	1.00, 1.40	0.049
0.0005		1.22	1.03, 1.44	0.019
0.005		1.22	1.04, 1.44	0.017
0.05		1.18	1.00, 1.39	0.051
0.5		1.22	1.04, 1.44	0.018
Ever smoking by 24 (1 cigarette or more)	2,931			
5×10^{-8}		1.25	1.16, 1.35	<0.001
0.0005		1.27	1.17, 1.38	<0.001
0.005		1.32	1.22, 1.43	<0.001
0.05		1.33	1.23, 1.44	<0.001
0.5		1.34	1.24, 1.44	<0.001
Smoking initiation (100 cigarettes or more) by 24	2,925			
5×10^{-8}		1.29	1.19, 1.39	<0.001
0.0005		1.38	1.27, 1.49	<0.001
0.005		1.46	1.34, 1.58	<0.001
0.05		1.49	1.37, 1.61	<0.001
0.5		1.49	1.37, 1.39	<0.001
Ever e-cigarette use by 24 among never smokers (<100 cigarettes*)	1,937			
5×10^{-8}		1.10	0.97, 1.26	0.150
0.0005		1.05	0.92, 1.20	0.464
0.005		1.12	0.98, 1.28	0.087

0.05		1.15	1.00, 1.31	0.046
0.5		1.18	1.04, 1.35	0.012
Ever e-cigarette use by 24 among never smokers (<1 cigarette)**	1037			
5×10^{-8}		1.28	0.94, 1.74	0.111
0.0005		1.10	0.81, 1.48	0.550
0.005		1.27	0.94, 1.71	0.124
0.05		1.39	1.02, 1.90	0.039
0.5		1.57	1.16, 2.12	0.004

Note: Ever smoking and smoking initiation models were included as positive controls. Analyses were adjusted for age, sex and principal components 1-10. *Never smokers in this analysis were defined as having smoked less than 100 cigarettes in their lifetime. **Never smokers in this analysis were defined as never having smoked a whole cigarette in their lifetime.

4.4.3 PRS for Smoking Initiation and Self-Reported E-cigarette Use

I observed positive associations between smoking initiation PRS and self-reported ever use of e-cigarettes by the age of 24 years ($OR_{10\text{-}8}$ = 1.24, 95% CI 1.14 to 1.34) and self-reported regular (at least once a month) e-cigarette use at 24 years ($OR_{10\text{-}8}$ = 1.18, 95% CI 1.00 to 1.40). I observed these associations at all *p*-value thresholds (Table 4.4). Among those who had never initiated smoking (i.e., smoked < 100 cigarettes in their lifetime), I found no clear evidence for an association between PRS for smoking initiation and ever e-cigarette use at the most stringent *p*-value thresholds. However, I found evidence of a positive association with PRS calculated with less stringent thresholds (p < 0.5 threshold OR = 1.18, 95% CI 1.04 to 1.35; Table 4.4). I found similar patterns of association among those who had never smoked any cigarettes (Table 4.4).

4.4.4 PRS for Smoking Initiation and Negative Controls

I observed a positive association between smoking initiation PRS and high number of sexual partners by 23 years ($OR_{10\text{-}8}$ = 1.15, 95% CI 1.05 to 1.26) and having ever gambled by 24 years ($OR_{10\text{-}8}$ = 1.12, 95% CI 1.03 to 1.22) at all *p*-value thresholds (Table 4.5). I found some evidence of a positive association between smoking initiation PRS and enjoying taking risks at 24 years ($OR_{0.005}$ = 1.11, 95% CI 1.03 to 1.19), but this was less clear at the more stringent thresholds (Table 4.5). I found no clear evidence of an

association between smoking initiation PRS and having been in trouble with the law since their 23rd birthday (Table 4.5).

Table 4.5. Associations between polygenic risk scores for smoking initiation with negative controls of risky behaviour.

Outcome *p*-value threshold	**n**	**OR**	**95% CI**	**p**
Number of sexual partners by 23 years*	2,505			
5×10^{-8}		1.15	1.05, 1.26	0.003
0.0005		1.12	1.02, 1.23	0.019
0.005		1.18	1.08, 1.29	<0.001
0.05		1.25	1.14, 1.37	<0.001
0.5		1.30	1.19, 1.43	<0.001
Been in trouble with the law since 23rd birthday	2,928			
5×10^{-8}		1.00	0.79, 1.28	0.988
0.0005		1.12	0.88, 1.43	0.352
0.005		1.11	0.87, 1.41	0.407
0.05		1.04	0.82, 1.33	0.745
0.5		0.90	0.71, 1.15	0.394
Enjoys taking risks at 24 years	2,932			
5×10^{-8}		1.06	0.98, 1.14	0.154
0.0005		1.05	0.98, 1.14	0.163
0.005		1.11	1.03, 1.19	0.005
0.05		1.09	1.01, 1.17	0.029
0.5		1.08	1.01, 1.16	0.033
Ever gambled by 24 years	2,899			
5×10^{-8}		1.12	1.03, 1.22	0.008
0.0005		1.16	1.07, 1.26	0.001
0.005		1.16	1.06, 1.26	0.001
0.05		1.20	1.10, 1.30	<0.001
0.5		1.15	1.06, 1.25	0.001

Note: Number of sexual partners, trouble with the law, enjoying risk-taking and gambling models were included as negative controls. Analyses were adjusted for age, sex and principal components 1-10. *Low (<11) vs. high (11 or more) number of sexual partners.

I found evidence of a positive association between smoking initiation PRS and hyperactivity at 7 years ($OR_{0.0005}$ = 1.10, 95% CI 1.04 to 1.16) but not at the most stringent threshold (Table 4.6). I also observed a positive association with CD at 7 years ($OR_{10\text{-}8}$ = 1.10, 95% CI 1.03 to 1.17) at all thresholds (Table 4.6). I found some evidence of a positive association between PRS and ODD specifically at the 0.0005 threshold ($OR_{0.0005}$

= 1.08, 95% CI 1.02 to 1.14). I also found a positive association with lower parental SEP ($OR_{10\text{-}8}$ = 1.08, 95% CI 1.01 to 1.16) at all thresholds (Table 4.7).

Table 4.6. Associations between polygenic risk scores for smoking initiation with negative controls of externalising disorders in childhood.

Outcome ***p*-value threshold**	**n**	**OR**	**95% CI**	**p**
Hyperactivity at 7 years	5,227			
5×10^{-8}		1.02	0.96, 1.08	0.511
0.0005		1.10	1.04, 1.16	0.001
0.005		1.14	1.08, 1.20	<0.001
0.05		1.14	1.08, 1.21	<0.001
0.5		1.15	1.08, 1.21	<0.001
Conduct disorder at 7 years	5,334			
5×10^{-8}		1.10	1.03, 1.17	0.004
0.0005		1.11	1.04, 1.19	0.001
0.005		1.11	1.04, 1.18	0.002
0.05		1.08	1.01, 1.15	0.021
0.5		1.08	1.01, 1.15	0.017
Oppositional defiant disorder at 7 years	5,325			
5×10^{-8}		1.02	0.96, 1.08	0.496
0.0005		1.08	1.02, 1.14	0.013
0.005		1.04	0.98, 1.10	0.200
0.05		1.04	0.98, 1.10	0.173
0.5		1.02	0.96, 1.08	0.529

Note: Hyperactivity and conduct disorder were assessed using the strengths and difficulties questionnaire (SDQ) and oppositional defiant disorder was assessed using the development and wellbeing assessment (DAWBA). All variables were recoded into binary outcomes (no disorder/symptoms versus borderline/disorder/symptoms).

Table 4.7. Associations between polygenic risk scores for smoking initiation with negative controls of socioeconomic indicators.

Outcome *p*-value threshold	**n**	**OR**	**95% CI**	**p**
Parental SEP at birth (manual)	6,702			
5×10^{-8}		1.08	1.01, 1.16	0.017
0.0005		1.13	1.06, 1.21	<0.001
0.005		1.16	1.09, 1.24	<0.001
0.05		1.11	1.03, 1.18	0.003
0.5		1.13	1.05, 1.20	<0.001

Note: SEP = Socioeconomic position. Parental SEP was based on the higher of the mother or partner's occupational social class using the Office of Population Censuses and Surveys (OPCS) occupation codes.

4.5 Discussion

In contrast to the results of Allegrini and colleagues (2019), when using the most recent GWAS for smoking initiation, PRS were strongly associated with ever e-cigarette use by 24 years. As expected, I observed an association of smoking initiation PRS and both ever smoking and smoking initiation. It was notable that the associations of the smoking initiation PRS and both smoking and e-cigarette use were of similar magnitude, although this could be due to the high correlation between the two behaviours.

The association between smoking initiation PRS and e-cigarette use could be explained by smoking causally influencing e-cigarette use. This hypothesis is supported by observational evidence; as discussed in Chapter 3, use of e-cigarettes for smoking cessation is common among both young adults in the UK (Khouja, Taylor, et al., 2020) and adults in Great Britain (Action on Smoking and Health, 2017). However, the associations observed among the restricted analysis and between the negative control outcomes suggest there may be other factors at play – there may be shared genetic risk factors that influence both behaviours. Among never smokers, I found weak evidence of an association between PRS for smoking initiation and e-cigarette use, which suggests that e-cigarette use is not simply caused by smoking but that there is a shared genetic aetiology influencing both behaviours. Hence, what appears to be a gateway between e-cigarette use and smoking in previous studies (described in Chapter 2), could actually be a shared genetic liability, and the order of use is coincidental or due to other factors such as perceived risk or misreporting of smoking status (Khouja, Munafò, et al., 2020).

Alternatively, the smoking initiation PRS may be capturing much more than just smoking or nicotine use. Using less stringent *p*-value thresholds to create the PRS increases the percentage variance that it explains of the phenotype, and therefore the power to detect an association (up to a point); using less stringent thresholds also increases the likelihood of capturing SNPs which are related to other factors than the exposure of interest, which adds noise and eventually results in less specific PRS that explain less

variance in the exposure and more variance in other (horizontally pleiotropic) correlated variables. Increasing magnitudes of association with PRS and negative controls at less stringent *p*-value thresholds suggests that the smoking initiation PRS is capturing, at least in part, a broad phenotype which is not entirely specific to smoking/nicotine.

Although I observed weaker associations between risk-taking factors and smoking initiation PRS compared to e-cigarette use and smoking, the associations are still relatively strong and consistent. Recent observational evidence also indicated a strong association between e-cigarette use and smoking prior to adjusting for risk-taking behaviours and other shared risk factors, but showed no clear evidence of an association after adjusting for risk-taking behaviours and other shared risk factors (Kim & Selya, 2019). This suggests that the association between smoking initiation PRS and both smoking initiation and e-cigarette use could be due to the PRS predicting general risk-taking behaviour, including both smoking and e-cigarette use.

I also found an association between the smoking initiation PRS and externalising disorders in childhood (7 years) which precedes the age at which cigarettes are first smoked in the vast majority of cases in this cohort (>99%), and therefore cannot be a causal effect of own smoking. However, this association could potentially be due to causal *in utero* effects of maternal smoking in pregnancy or maternal smoking in childhood, since maternal and offspring genotype will be correlated. ADHD has been shown to be associated with in utero smoke exposure, and the effects are stronger among genetically-related mother-offspring pairs compared to unrelated mother-offspring pairs, suggesting an inherited effect (Thapar et al., 2009). Nevertheless, combined with evidence that liability to ADHD increases the likelihood of smoking initiation and vice versa (Treur et al., 2019), our results suggest the possibility that the smoking initiation PRS is capturing a broad impulsivity phenotype.

The association observed between smoking initiation PRS and parental SEP also suggests the PRS is capturing a broader phenotype than one relating specifically to nicotine or smoking. It is impossible for parental SEP at birth to be caused by the young adults' own smoking, so I can be confident that this association is not a result of own smoking causing SEP. This suggests that the smoking initiation PRS could actually be capturing an even broader phenotype, encompassing sociodemographic factors as well as risk-taking

behaviour. Although introducing noise is to be expected when using the less stringent *p*-value thresholds for SNP inclusion, it is concerning that the PRS created using the genome-wide significant threshold appear to be capturing a broader phenotype than the one intended as these PRS may be used in Mendelian randomisation (MR) analysis.

The associations observed here have important implications for the use of smoking initiation PRS in MR. This method is often implemented to provide unconfounded causal estimates, as long as the assumptions of MR hold (Davey Smith & Ebrahim, 2003). One assumption is that the genetic instrument (e.g., smoking initiation PRS) is not associated with any confounders (e.g., risk-taking, childhood externalising disorders, SEP). The association we observed between smoking initiation PRS and negative control outcomes, even when restricted to only genome-wide significant SNPs, indicates that smoking initiation PRS may not be a valid instrument to use in MR to investigate the causal effects of smoking initiation. This emphasises the importance of using pleiotropy robust methods (e.g., MR-Egger). Another assumption is the InSIDE (Instrument Strength Independent of Direct Effect) assumption that SNP-exposure effects (e.g., the effect of smoking initiation SNPs on smoking initiation) should not be correlated with horizontal pleiotropic effects (e.g., the effect of smoking initiation SNPs on broad risk-taking behaviour). The association observed between the smoking initiation PRS and multiple risk-taking behaviours and externalising disorders in childhood suggests that the smoking initiation SNPs may be capturing a broader phenotype, such as risk-taking, which is not specific to smoking or nicotine, and thus this assumption may be violated. One approach which could be used to address this is Steiger filtering which can be used to exclude SNPs which explain more variance in the *outcome* over and above the variance in the *exposure* (Hemani et al., 2018; Hemani et al., 2017). The same approach can be applied in MR studies using smoking initiation PRS to remove SNPs which explain more variance in the negative control outcomes used in this study (or other phenotypes/proxies for risk-taking behaviour) than variance in smoking initiation. However, if the InSIDE assumption is perfectly violated (i.e., if the SNP effect on broad risk-taking causes smoking initiation), the smoking initiation PRS will be an invalid instrument using any MR method. At the very least, triangulating evidence across multiple MR methods (e.g., median weighted and mode based) would be advised in MR studies using smoking initiation PRS but, ideally, other causal inference methods should also be used. Additionally, triangulating evidence using other measures of smoking

which are not as highly associated with impulsivity, such as smoking heaviness, could be useful if including such measures would be appropriate for the research question.

There are a number of limitations of this study. First, the relatively low sample size – particularly when investigating associations with regular e-cigarette use and restricting to never smokers. The results of my restricted analysis may suffer from low power as a result. Similarly, lack of power may explain why I found no clear association between the smoking initiation PRS and being in trouble with the law; only 7% of the study sample had been in trouble with the law in the past year. Second, restricting my analysis to never smokers could have introduced collider bias (Cole et al., 2010). Collider bias occurs when you condition (restrict, stratify or adjust regression analyses) on, or control for a collider (a variable which is influenced by both the exposure and the outcome) in your analysis. In this analysis I found that smoking initiation PRS were strongly associated with smoking initiation; if e-cigarette use *also* causes young adults to smoke (as discussed in Chapter 2), then smoking status is a collider and conditioning on this variable (i.e., restricting analysis to never smokers) may inflate any association between smoking initiation PRS and e-cigarette use. In this sample, there are very few individuals for whom e-cigarette use preceded smoking, so it is unlikely that collider bias has been introduced in this way. However, if both smoking initiation PRS and unmeasured confounders influence smoking initiation (as shown in Figure 4.2), then smoking initiation is a collider and conditioning on smoking initiation (i.e., restricting the analysis to never smokers) could induce bias (Paternoster, Tilling, & Davey Smith, 2017).

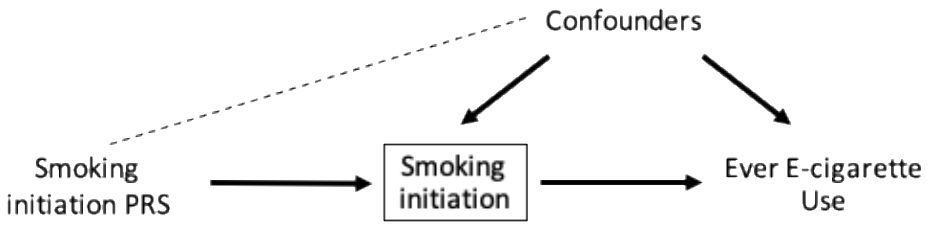


Figure 4.2. A directed acyclic graph demonstrating potential collider bias introduced by conditioning on a collider (smoking initiation).

Third, although my restricted analysis (excluding smokers) could be used to explore the gateway hypothesis, this cohort is not appropriate to directly study the gateway hypothesis. The young adults in ALSPAC were approximately 17 years old when e-cigarettes became widely available, and therefore were exposed to cigarettes earlier in their adolescence than e-cigarettes and had more opportunity to smoke than use e-cigarettes than later birth cohorts. Furthermore, I found in Chapter 3 that using e-cigarettes for smoking cessation/reduction is common, so given the age and nature of the sample as well as the reasons provided for e-cigarette use, it is not unexpected that we would observe an association between a PRS that predicts smoking initiation and e-cigarette use. However, it is unexpected that we would observe an association with risk-taking, which suggests a potential broader risk-taking behaviour underlying this phenotype. Future research should explore this association in a larger sample of individuals with exposure to both cigarettes and e-cigarettes during adolescence. Fourth, the attrition rate in ALSPAC is considerable – only 2,905 of the 7,859 non-related participants of European ancestry with genetic data responded to the questions about vaping in the 24-year questionnaire – and missingness in this cohort has previously been associated with smoking initiation PRS (Taylor et al., 2018). I found that those with higher smoking initiation PRS were less likely to have been included in this analysis due to attrition ($OR_{10\text{-}8}$ per standard deviation of smoking initiation PRS = 0.87, 95% CI 0.83 to 0.91) so my estimates may be biased by selection and the association *could* be stronger than observed here. However, interpretation of any study including smoking initiation PRS will be difficult as the association between smoking initiation PRS and attrition could induce bias such as collider bias (Munafò, Tilling, Taylor, Evans, & Davey Smith, 2018).

4.6 Chapter Summary

In conclusion, I found some evidence to suggest there is a shared genetic aetiology between smoking and e-cigarette use. This supports my hypothesis in Chapter 2; what appears to be a gateway between e-cigarette use and smoking in observational studies could actually be a result of shared liability, specifically a shared genetic aetiology. However, this does not rule out a potential gateway effect, and there is likely a causal relationship between smoking and later e-cigarette use given the reasons provided in Chapter 3. Additionally, the smoking initiation PRS is associated with risky behaviour,

externalising disorders in childhood and parental SEP at birth. This suggests the smoking initiation PRS is not specific to smoking or nicotine use but is capturing something much broader, potentially a risk-taking phenotype. Therefore, adolescents and young adults may use e-cigarettes because they are genetically predisposed to risk-taking and they consider e-cigarette use to be risky. This could explain why researchers have observed associations between e-cigarette use and subsequent smoking in observational studies (as discussed in Chapter 2) as smoking is also considered to be a risky behaviour. Future research is needed to explore this further in a population which has been exposed to both e-cigarettes and cigarettes in adolescence. Researchers should also be cautious about inferring causality from MR analyses using smoking initiation PRS. In Chapter 5, I use MR methods to explore the potential health impact of using e-cigarettes.

Chapter 5 What is the health impact of nicotine exposure versus other constituents of smoking?

This chapter closely resembles sections from the following pre-print:

Khouja, J. N., Sanderson, E., Wootton, R. E., Taylor, A. E., & Munafò, M. R. (2021). A multivariable Mendelian randomisation study exploring the direct effects of nicotine on health compared with the other constituents of tobacco smoke: Implications for e-cigarette use. MedRxiv; doi: https://doi.org/10.1101/2021.01.12.21249493.

I took the lead in designing, implementing and interpreting this study with the support of my co-authors. Specifically, MM, RW, AT and ES provided feedback on the design of the study and during the preparation of the manuscript. I was responsible for the data analysis and was responsible for writing the first draft of the manuscript and editing the manuscript in response to comments from co-authors and reviewers.

5.1 Chapter Overview

In Chapter 3 I found that vaping for reasons related to smoking cessation is common among young adults in the UK, but the long-term health effects of vaping (such as the long-term effects of nicotine use without tobacco smoke exposure) are currently unknown. As e-cigarettes have not been available for long enough to observe the long-term health consequences of use, the methods used to determine the effects of smoking (i.e., longitudinal observational methods) cannot be adopted to explore the effects of vaping. Furthermore, it is difficult to disentangle the effects of e-cigarette use from the effects of smoking given they are highly correlated. In this Chapter, I use a novel method (multivariable Mendelian randomisation [MVMR]) to explore the effects of using nicotine – the addictive common constituent of cigarettes and e-cigarettes – while accounting for the effects of other constituents of cigarette smoke.

5.2 Introduction

Of an estimated 3.6 million e-cigarette users in Great Britain, 22% use e-cigarettes to aid quitting (Action on Smoking and Health, 2019a) and in Chapter 3, I found that vaping to quit smoking was the second most common reason given for using e-cigarettes. Although evidence suggests that e-cigarettes may reduce harm by aiding smoking cessation (Hartmann-Boyce et al., 2020; McNeill et al., 2015), the long-term health effects of nicotine exposure via e-cigarette use remain unknown.

In contrast, the long-term health outcomes of smoking are well-known due to the abundance of observational evidence demonstrating associations between smoking and health issues such as chronic obstructive pulmonary disease (COPD), coronary heart disease (CHD) and poor lung function (Das, 2003; Jayes et al., 2016; Khan, Shabbir, Ansari, & Zia, 2010; West, 2017). Consistent evidence across many observational studies provides strong support for a causal effect (Hill, 1965), which is further supported by genome-wide association studies that identify smoking-related genetic variants when examining these outcomes (Gage, Davey Smith, Ware, Flint, & Munafo, 2016; Hobbs et al., 2017; Polfus et al., 2013; Zhang et al., 2014). However, it remains unclear which constituents of tobacco smoke (e.g., nicotine, carbon monoxide etc.) negatively impact health, or have the largest effects.

Until e-cigarettes became widely available in 2007, nicotine replacement therapy (NRT) was the primary source of nicotine without tobacco. Long-term NRT use is rare among ex-smokers (Shahab, Dobbie, et al., 2017) and non-smokers (Jackson et al., 2019); consequently, there is little evidence of the long-term effects of nicotine use when not consumed in tobacco products. Ideally, to explore the effect of long-term nicotine use on health outcomes, I would conduct a randomised controlled trial (RCT) with never-smokers randomised to use nicotine or placebo and observe the differences between groups for a range of health outcomes over time. However, it would be unethical to expose never-smokers to an addictive and potentially harmful drug, and the individuals would need to be followed up for decades to observe the long-term effects.

As I described in Chapter 1, Mendelian randomisation (MR) is a method which is often used to infer causality, particularly where an RCT would be unethical or impossible

(Davey Smith & Ebrahim, 2003). The method assumes that the laws of Mendelian genetics (segregation and independent assortment) are held at a population level i.e., a random assortment of genes are transferred from parents to their offspring (Nitsch et al., 2006). For example, if an individual inherits the rs16969968 genetic variant which predisposes them to be more tolerant of nicotine, then they are likely to be a heavier smoker (i.e., smoke one more cigarette per day per risk allele) on average than an individual with no risk alleles (Ware, van den Bree, & Munafò, 2011). In theory, the inheritance of these genetic variants is independent of confounding factors which often distort observational evidence, and therefore mimics the randomisation process in an RCT, reducing issues of both confounding and reverse causality (Nitsch et al., 2006; Swanson et al., 2017). The MR method estimates the total causal effect of one exposure on one outcome. For example, to explore the potential harm of using nicotine-containing products (e.g., e-cigarettes), I could use MR methods to estimate the total effect of e-cigarette use on COPD. However, to conduct MR analysis we require large genome-wide association studies (GWAS) of the exposure, to identify genetic variants that can be used as proxies for this exposure. At present, there are no published GWAS of e-cigarette use, nor are there currently any cohorts or consortia with sufficient numbers of e-cigarette users to conduct this analysis. Furthermore, e-cigarette use and smoking are highly correlated (Action on Smoking and Health, 2019a) and may share a genetic aetiology (as found in Chapter 4); to ensure any associations found are not due to confounding effects of smoking, the GWAS should be restricted to never-smokers, but as I discussed in Chapter 2, few never-smokers regularly vape (Action on Smoking and Health, 2019a; Bauld et al., 2017; Hammond et al., 2019).

As I described in Chapter 1, multivariable MR (MVMR) is an extension of the inverse variance weighted (IVW) MR method; rather than calculating the total effect, MVMR is used to explore the direct causal effect of two or more exposures on an outcome (Lawlor et al., 2019; Sanderson et al., 2019). When two exposures are related, MVMR can estimate the effect of one exposure on an outcome while accounting for the effect of the other exposure on the outcome (i.e., the direct effect) even when there is overlap in the genetic effects on the two exposures. As cigarettes contain nicotine, smoke exposure and nicotine intake are highly related; therefore, MVMR is a suitable method to explore the direct effects of nicotine versus the direct effect of all of the other constituents of tobacco smoke on smoking-related health outcomes.

GWAS have previously identified genetic variants associated with smoking heaviness (i.e., average number of cigarettes smoked per day [CPD]) as well as cotinine – a highly specific biomarker which captures recent nicotine exposure (Figure 5.1). Given that cotinine has a longer half-life (16 hours) than nicotine (2 hours), and that 70-80% of nicotine is rapidly metabolised into cotinine, cotinine levels are a good proxy for nicotine exposure (Benowitz et al., 2009).

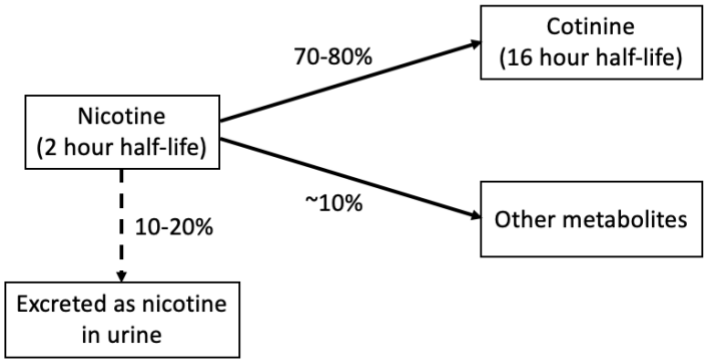


Figure 5.1. A diagram to show the by-products of nicotine metabolism. Adapted from Benowitz et al. (2009).

By using these genetic variants as proxies for nicotine (G_N) and smoking (G_S) – including those variants which predict both nicotine and smoking heaviness (G_{SN}) – in an MVMR analysis (Figure 5.2a), I can explore the direct effect of cotinine while taking into account the effect of smoke exposure (Figure 2b) and vice versa (Figure 5.2c). The total effects of smoking heaviness on health outcomes includes the effect of cotinine on health outcomes, but by exploring the direct effect of smoking heaviness while controlling for the direct effect of cotinine (and confounding factors), I can observe the effect of the remaining constituents of tobacco smoke aside from cotinine (Figure 5.2c). In other words, among smokers, I can explore the health effects caused by nicotine versus the health effects caused by the other constituents of tobacco smoke exposure (aside from nicotine).

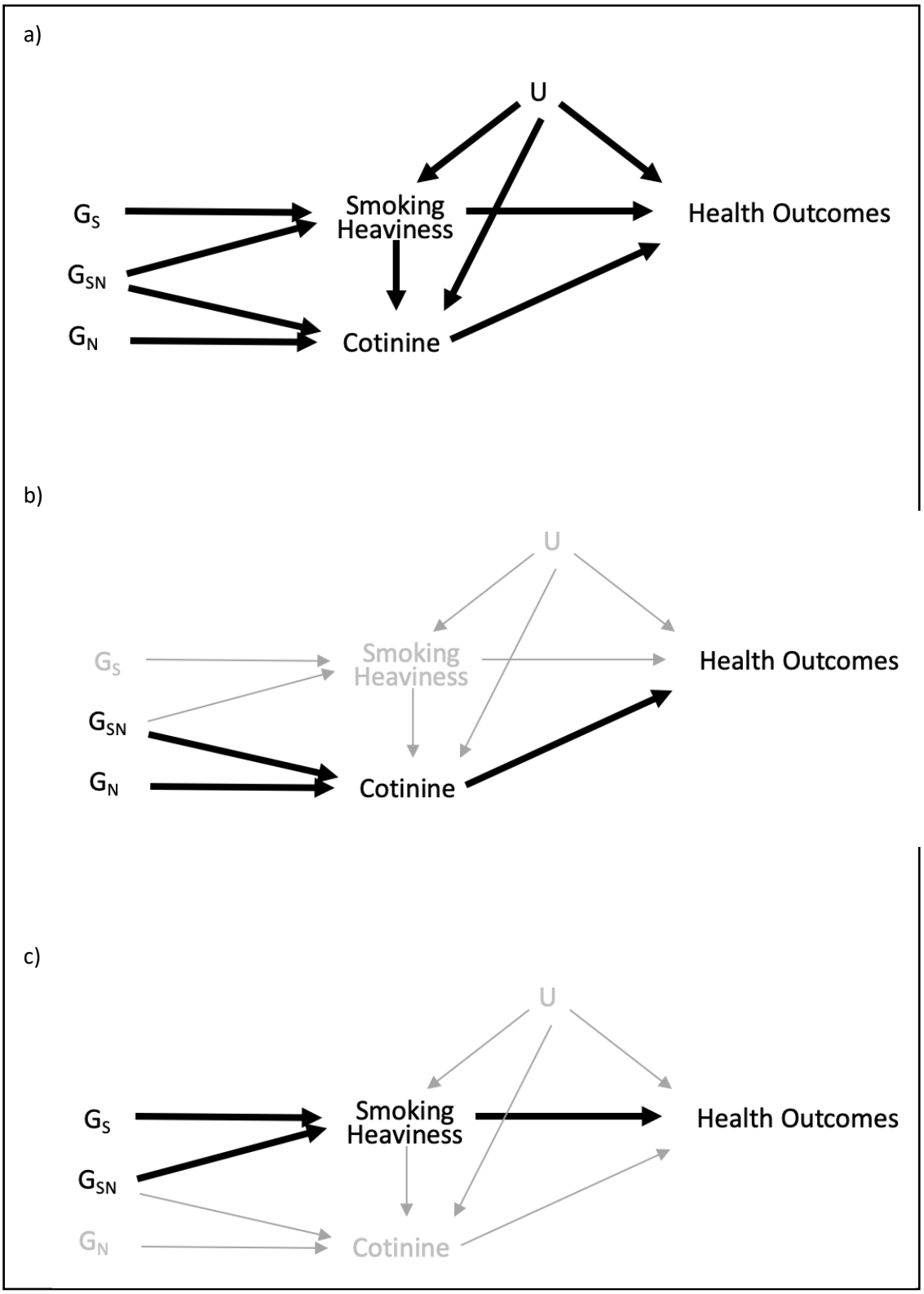


Figure 5.2. Directed acyclic graphs to show the relationship between genetic instruments (G_S, G_{SN}, and G_N), exposures (smoking heaviness and cotinine), confounding (U) and outcomes (health outcomes) in a multivariable Mendelian randomisation analysis.

The aim of this study was therefore to explore the direct effects of nicotine (as measured by cotinine, independent of cigarettes per day) compared with the direct effect of all other constituents of tobacco smoke aside from nicotine (as measured by CPD, independent of cotinine) on health outcomes known to be caused by smoking.

5.3 Methods

5.3.1 Data Sources

The data sources for the exposures and outcomes (including sample sizes) are shown in Figure 5.3. I describe each of the data sets obtained from these sources as either individual-level or summary-level data. As described in Chapter 1, individual-level data consist of genetic, exposure and outcome data for all individual participants with which genetic associations can be calculated. Where individual-level data are provided, summary-level data can be generated for further analysis (Figure 5.3). Summary-level data contain only the overall genetic associations with the exposure and outcome for the whole sample, and can be used to identify suitable genetic instruments and the effect sizes of the instrument-phenotype association for inclusion in MR and MVMR analysis.

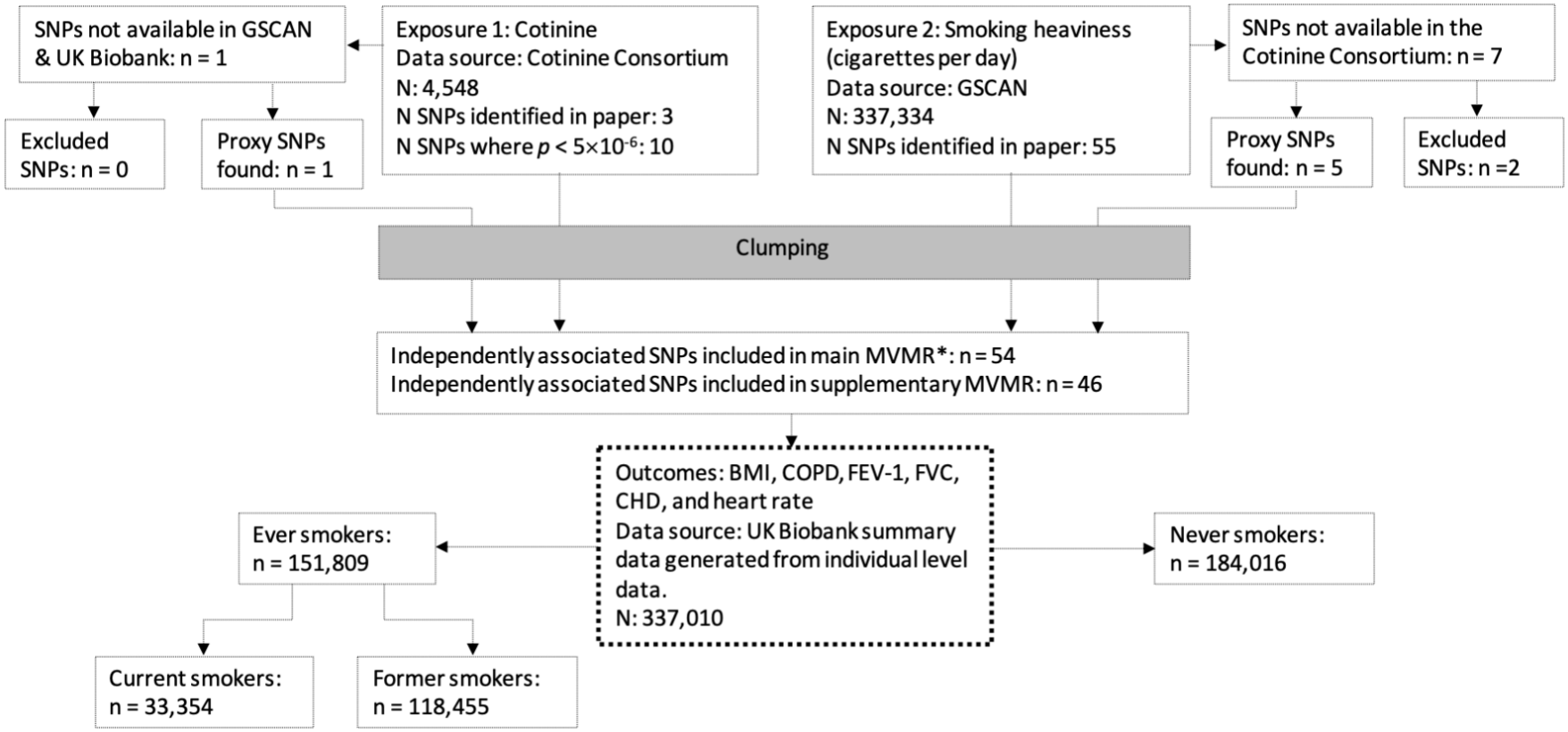


Figure 5.3. A flow chart describing the SNP inclusion process for the Multivariable Mendelian Randomisation analyses.

Note: SNPs = single nucleotide polymorphisms; BMI = body mass index; COPD = chronic obstructive pulmonary disease; FEV-1 = forced expiratory volume in 1 second; FVC = forced vital capacity; CHD = chronic heart disease. *SNPs were selected for inclusion in the main analysis using a *p*-value threshold ($p < 5\times10^{-6}$); ** SNPs were selected for inclusion in the supplementary analysis was completed using the genome wide significant threshold ($p < 5\times10^{-8}$). Boxes with dashed borders indicate where individual-level data was used to generate summary-level data. Boxes with solid borders indicate summary-level data. 1,185 UK Biobank participants were excluded who preferred not to state their smoking status.

5.3.1.1 The Cotinine Consortium. Ware et al. (2016) report summary level statistics from a GWAS meta-analysis of cotinine levels (per standard deviation change) among daily smokers of European ancestry. The meta-analysis is comprised of data from 11 studies in the Cotinine Consortium (N = 4,548). Cotinine levels were provided by daily smokers at or after 17 years of age and were determined from serum, urine, or plasma samples (dependent on the sample available) and quantified using immunoassay, radioimmunoassay or mass spectrometry. Specific cotinine thresholds (dependent on the assessment method) were used to reduce the inclusion of non-smokers and non-daily smokers in the analysis. Further information, including the age of participants, location, setting and assessment method for each cohort are detailed in the supplementary material of Ware et al. (2016). Cotinine levels were standardised (i.e., transformed into z-scores) at the individual study-level prior to conducting the GWAS, therefore the GWAS were conducted using the standardised cotinine levels. SNPs were reported as independent if they reached genome-wide significance using an iterative process of conditional analyses. This process involved re-running the meta-analysis for each region identified by the initial meta-analysis while conditioning on the SNP with the lowest *p*-value. The next strongest signal was identified from this conditional meta-analysis. This SNP was then conditioned on in the second conditional meta-analysis. This process was repeated until no residual signal remained below a threshold of $p < 5\times10^{-8}$.

5.3.1.2 GSCAN. The GWAS & Sequencing Consortium of Alcohol and Nicotine use (GSCAN) report summary level statistics from a GWAS of smoking heaviness (Liu et al., 2019). This GWAS was based on 337,334 participants of European ancestry from 29 cohorts. Summary statistics, with the UK Biobank (N = 120,744) removed, were obtained through correspondence with GSCAN to eliminate data overlap with the target sample. Due to data sharing restrictions, 23andMe were also excluded from this data (N = 73,380) leaving a total sample size of 143,210. Smoking heaviness was defined as the number of CPD smoked either as a current or former smoker (including hand-rolled and manufactured cigarettes). Where the number of cigarettes per day was reported as an integer, responses were binned into 5 groups (1-5; 6-15; 16-25; 26-35; 36+ cigarettes per day); where studies had created pre-defined bins, these were used. For the majority of studies, a single question was used to determine the number of CPD (e.g., how many cigarettes do/did you smoke per day?). CPD was standardised using the standard

deviation (calculated from the weighted average prevalence across all studies). SNPs were reported as independent if they explained additional variance in conditional analyses using a partial correlation-based score statistic (Jiang et al., 2018).

5.3.1.3 UK Biobank. I obtained individual-level data from UK Biobank, a population-based health research resource consisting of approximately 500,000 people, aged between 38 years and 73 years, who were recruited between the years 2006 and 2010 from across the UK (Allen, Sudlow, Peakman, Collins, & Biobank, 2014). With a particular focus on identifying determinants of human diseases in middle-aged and older individuals, participants provided a wide range of health information (data available at www.ukbiobank.ac.uk). A full description of the study design, participants and quality control (QC) methods have been described in detail previously (Allen et al., 2014; Mitchell et al., 2019). UK Biobank received ethics approval from the Research Ethics Committee (REC reference for UK Biobank is 11/NW/0382; project number 9142).

5.3.2 Health Outcomes

I have provided UK Biobank field IDs for each health outcome. IDs can be entered into the online variable search platform where information can be found on measurement (available at http://biobank.ndph.ox.ac.uk/showcase/search.cgi).

5.3.2.1 Body Mass Index. Standing height and weight were measured during a UK Biobank Assessment Centre visit (field IDs 50 and 21002). BMI was calculated as: weight (kg)/height $(m)^2$.

5.3.2.2 Chronic Obstructive Pulmonary Disease. COPD cases were identified as participants who responded that they had been diagnosed with COPD in response to the question "Has a doctor ever told you that you have had any of the conditions below?" (field ID 22130).

5.3.2.3 Forced Expiratory Volume and Forced Vital Capacity. Forced expiratory volume in 1 second (FEV-1) and forced vital capacity (FVC) were measured multiple times using a Vitalograph spirometer (field IDs 20150 and 20151). I used the 'best measure' of both FEV-1 and FVC which were identified as the highest value recorded with no measurement issues reported.

5.3.2.4 Coronary Heart Disease. CHD diagnosis was determined using linked hospital admission data (field IDs 41270 and 41271) with ICD codes relating to ischemic heart disease (ICD-9 410-414; ICD-9 4100-4149; ICD-10 I20-I25). The measure included angina pectoris, acute myocardial infarction, subsequent myocardial infarction, certain current complications following acute myocardial infarction, other acute ischaemic heart diseases, and chronic ischaemic heart disease.

5.3.2.5 Heart Rate. Heart rate (beats per minute [bpm]) was assessed on multiple occasions per session (field ID 102). Heart rate can be affected by numerous factors such as exercise (Evans, 1985) and stress (Kim, Cheon, Bai, Lee, & Koo, 2018). To allow time for the participant's heart rate to normalise during the session, I used the second measure taken within the session.

5.3.2.6 Smoking Status. In UK Biobank, smoking status was categorised as never, previous and current smoking (field 20116). From this variable, I derived an 'ever smoking' variable which was defined as currently or having previously smoked occasionally, most days or daily (i.e., having smoked more than just once or twice). Current smoking was defined as currently smoking occasionally, most days or daily. Former smoking (referred to as 'previous smoking' in UK Biobank) was defined as not currently smoking but having previously smoked occasionally, most days or daily (i.e., more than just once or twice). Those who have tried smoking once or twice or who have never smoked were categorised as never smokers.

5.3.3 Statistical Analysis

After restricting the sample to individuals of White British ancestry (Bycroft et al., 2018) and excluding those with mismatched sex, with missing array data, who were related or withdrew their consent to participate, the sample size was 337,010 (Mitchell et al., 2019). Using individual-level data from UK Biobank, I generated summary-level data by regressing each SNP on each of the health outcomes, adjusting for 10 principal components of population stratification. As shown in Figure 5.3, four datasets were generated according to smoking status: ever smokers (i.e., both current and former smokers; n = 151,809), current smokers (n = 33,354), former smokers (n = 118,455), and never smokers (n = 184,016). Binary outcomes were estimated using logistic regressions, and continuous outcomes using linear regressions.

All analyses were carried out in Stata 15.1 (StataCorp, 2017).

5.3.4.1 Selection of genetic variants. Genetic variants related to the phenotype of interest (cotinine levels or smoking heaviness) were selected for inclusion in the analysis based on the reported results of the relevant GWAS (Figure 5.3). Independent SNPs at the genome-wide significant level ($p < 5\times10^{-8}$) were selected for inclusion – 55 SNPs were identified as associated with smoking heaviness independently of any other SNP associations (Liu et al., 2019) and 3 SNPs were identified as associated with cotinine levels independently of any other SNP associations (Ware et al., 2016).

For MVMR analyses, all included SNPs (i.e., those relating to smoking heaviness as well as those relating to cotinine levels) must also be independent of each other, so an additional clumping stage was added to ensure overall SNP independence (LD $R^2 < 0.1$, > 500 kb). As I used a different method of determining independence in this study (clumping) compared to the GWAS of cotinine levels (conditional independence), 1 SNP which was considered conditionally independently associated with cotinine levels in the original GWAS (rs57064725) was not considered independent from the other cotinine SNPs in the MVMR analysis. As fewer SNPs were associated with cotinine than with smoking heaviness, where a SNP associated with cotinine levels was in LD with a SNP associated with smoking heaviness, the SNP associated with smoking heaviness was removed.

Due to the limited number of independent SNPs associated with cotinine levels at the genome-wide significant threshold (n = 2), I lowered the significance threshold used for inclusion of cotinine SNPs in the main analyses to $p < 5x10^{-6}$. Where a SNP was identified for inclusion but was not available in either of the other summary data sets (e.g., available in the Cotinine Consortium summary data but not available in the GSCAN summary data or Biobank data), I selected proxy SNPs with a minimum linkage disequilibrium (LD) R^2 of 0.8. Proxy SNPs were available for 5 SNPs which were associated with smoking heaviness but were not available in the cotinine GWAS summary data, and for 1 SNP which was associated with cotinine but not available in the smoking heaviness GWAS summary data or UK Biobank data. I excluded 2 SNPs which were associated with smoking heaviness which were not available in the cotinine GWAS summary data and for which no suitable proxies were available (i.e., no available SNPs

with LD $R^2 > 0.8$). Details of the SNPs included in each analysis, and proxies used, are provided in Appendix 8.

Instrument strength was tested using the conditional F-statistic for MVMR and Cochran *Q* statistic (Sanderson, Spiller, & Bowden, 2020). Instrument strength is quantified using the mean F-statistic in the IVW method (Bowden & Holmes, 2019). As a general rule, the F-statistic should be greater than 10. The Cochran *Q* statistic tests for instrument strength in the two sample summary data setting by detecting heterogeneity among causal estimates (Bowden & Holmes, 2019). *Q* estimates should be less than the number of SNPs included in the model.

5.3.4.2 Univariable Mendelian randomisation. For comparison with the main analysis, I considered the total effect of both cotinine levels and smoking heaviness on each health outcome (BMI, COPD, FEV-1, FVC, CHD and heart rate) using MR with summary data from the Cotinine Consortium and GSCAN, and summary data (generated using individual-level data) from UK Biobank. I used four complimentary methods (inverse variance weighted [IVW], MR-Egger, weighted median based estimation and weighted modal based estimation) (Bowden et al., 2015; Bowden, Davey Smith, et al., 2016; Burgess, Butterworth, & Thompson, 2013; Hartwig, Davey Smith, & Bowden, 2017). Using a variety of MR methods with different assumptions with respect to horizontal pleiotropy – which occurs when single genetic variants influence multiple phenotypes (Davey Smith & Hemani, 2014) – facilitates understanding of whether the effect of the exposure on the outcome is causal. Consistent results across these methods provide stronger evidence to support a true causal effect which is not the result of a false positive (Lawlor, Tilling, & Davey Smith, 2016). I also estimated the weighted regression dilution (I^2_{GX}) for each MR-Egger analysis (Bowden, Del Greco, et al., 2016). I applied simulation extrapolation SIMEX (Lederer & Küchenhoff, 2006) corrections to MR-Egger analysis where I^2_{GX} estimates were below 0.9 (which would indicate the effect estimate is biased by 10% due to measurement error) (Bowden, Del Greco, et al., 2016). This analysis was first restricted to ever smokers to capture the long-term effects of smoking/nicotine use (i.e., including ex-smokers who may have quit smoking due to developing a health issue), and then further restricted to current smokers only to explore the potential effects of smoking cessation (i.e., recoverable effects). If a poor health outcome was found among current smokers but not ever smokers, it would

indicate that that health outcome is short term and may improve following smoking cessation.

5.3.4.3 Multivariable Mendelian randomisation. I explored the direct effects of cotinine levels and smoking heaviness individually (i.e., accounting for the other phenotype) using MVMR with two complimentary methods – IVW and MVMR-Egger (Sanderson et al., 2019). All of the SNPs included in this analysis were associated with either cotinine levels or smoking heaviness (or both). To explore the recoverable and long-term outcomes of smoking, this analysis was first restricted to ever smokers, and then restricted to current smokers.

I also stratified the analysis by former smoking status to further explore recoverable effects, and I additionally stratified the analysis by ever smoking status to explore potential horizontal pleiotropy. Horizontal pleiotropy is problematic for the interpretation of MR results because horizontally pleiotropic genetic variants are not valid instruments – effects observed among never smokers could indicate horizontally pleiotropic effects (i.e., the included SNPs influence the outcome directly, or via another phenotype, but not through smoking), or it could indicate misreporting of smoking status or population stratification.

5.4 Results

5.4.1 Descriptive Statistics

Of the 337,010 individuals with available data in UK Biobank, 54% were male, the mean age was 56, and the average BMI was 27.39. A total of 1,245 (1%) had a diagnosis of COPD and 28,652 (9%) had a diagnosis of CHD. Average FEV-1 was 2.87 litres, average FVC was 3.80 litres, and average heart rate was 68.98 bpm. A total of 184,016 (55%) were never smokers and 151,809 (45%) were ever smokers, of whom 33,354 (10%) were current smokers and 118,455 (35%) were former smokers. A total of 1,185 UK Biobank participants were excluded who preferred not to state their smoking status.

5.4.2 Univariable Mendelian Randomisation

The results of the univariable MR analysis of cotinine levels (n = 10 SNPs) and smoking-related health outcomes among ever smokers and current smokers are displayed in Table 5.1. The results of the MR analysis of smoking heaviness (cigarettes per day; n = 54 SNPs) and smoking-related health outcomes among ever smokers and current smokers are shown in Table 5.2. Results are reported per standard deviation (SD) increase in the exposure phenotype (i.e., cotinine/cigarettes per day) and relate to the results of the IVW estimation except where otherwise specified. Tests of the weighted regression dilution (I^2_{GX}), instrument validity and heterogeneity can be found in Appendix 9 and Appendix 10.

Table 5.1. Univariable Mendelian randomisation analysis of cotinine and smoking-related health outcomes among ever and current smokers (n = 10 SNPs).

Health outcome	MR method	Ever smokers			Current smokers		
		Effect	95% CI	*p*-value	Effect	95% CI	*p*-value
BMI	IVW	0.03	-0.11, 0.16	0.671	-0.35	-0.75, 0.05	0.086
	MR-Egger (SIMEX)	0.03	-0.91, 0.98	0.944	-1.73	-4.07, 0.61	0.147
	W. median	0.01	-0.12, 0.14	0.934	-0.39	-0.77, -0.02	0.041
	W. mode	-0.01	-0.21, 0.19	0.925	0.22	-0.60, 1.05	0.593
COPD	IVW	0.24	-0.12, 0.60	0.189	0.53	-0.10, 1.15	0.099
	MR-Egger (SIMEX)	0.92	-1.75, 3.58	0.502	1.07	-3.32, 5.47	0.633
	W. median	0.03	-0.30, 0.36	0.842	0.54	-0.11, 1.19	0.102
	W. mode	-0.01	-0.46, 0.44	0.968	0.44	-0.65, 1.53	0.430
FEV-1	IVW	-0.03	-0.06, 0.00	0.041	-0.06	-0.11, -0.02	0.005
	MR-Egger (SIMEX)	-0.17	-0.31, -0.02	0.023	-0.23	-0.51, 0.06	0.118
	W. median	-0.03	-0.06, 0.00	0.021	-0.05	-0.10, 0.01	0.100
	W. mode	-0.03	-0.08, 0.01	0.150	-0.01	-0.10, 0.07	0.736
FVC	IVW	-0.02	-0.04, 0.00	0.049	-0.05	-0.10, -0.01	0.024
	MR-Egger (SIMEX)	-0.10	-0.20, 0.00	0.053	-0.12	-0.27, 0.02	0.103
	W. median	-0.02	-0.05, 0.00	0.100	-0.04	-0.10, 0.02	0.205
	W. mode	-0.04	-0.08, 0.00	0.060	-0.03	-0.11, 0.05	0.487
CHD	IVW	-0.01	-0.10, 0.09	0.860	0.00	-0.15, 0.15	0.982
	MR-Egger (SIMEX)	0.03	-0.33, 0.39	0.874	-0.04	-0.82, 0.74	0.919
	W. median	0.05	-0.03, 0.14	0.244	0.09	-0.08, 0.26	0.292
	W. mode	0.07	-0.03, 0.18	0.172	0.07	-0.15, 0.30	0.529
Heart Rate	IVW	0.27	-0.02, 0.56	0.070	0.75	0.10, 1.39	0.023
	MR-Egger (SIMEX)	0.98	-0.74, 2.70	0.262	3.24	-0.29, 6.77	0.072
	W. median	0.23	-0.10, 0.57	0.172	0.46	-0.26, 1.18	0.207
	W. mode	0.18	-0.45, 0.82	0.575	0.12	-1.14, 1.38	0.850

Note: A *p*-threshold of 5×10^{-6} was used to determine the single nucleotide polymorphisms (SNPs) associated with cotinine. BMI = body mass index; COPD = chronic obstructive pulmonary disease; FEV-1 = forced expiratory volume in 1 second; FVC = forced vital capacity; CHD = chronic heart disease; IVW = inverse variance weighted; w. median = weighted median; w. mode = weighted mode. Effects are betas for continuous variables (BMI, FEV-1, FVC and HR) and log odds ratios for binary outcomes (COPD and CHD) per standard deviation increase in cotinine levels.

5.4.2.1 Total effects of cotinine levels on health outcomes. The total effects of cotinine levels on health outcomes among ever smokers are displayed in Table 5.1. Among ever smokers, there was some evidence to suggest that higher cotinine levels cause decreased FEV-1 (β = -0.03 litres, 95% CI -0.06 to 0.00 per SD increase in cotinine levels) and FVC (β = -0.02 litres, 95% CI -0.04 to 0.00 per SD increase in cotinine levels) and weak evidence to suggest that higher cotinine levels increased heart rate (β = 0.27 bpm, 95% CI -0.02 to 0.56 per SD increase in cotinine levels). These results were consistently in the same direction for all methods. There was no clear evidence to suggest a causal effect of cotinine levels on BMI or the risk of COPD or CHD.

The total effects of cotinine levels on health outcomes among current smokers are displayed in Table 5.1. Among current smokers, there was evidence to suggest that higher cotinine levels cause decreased BMI (weighted median β = -0.39 kg/m^2, 95% CI -0.77 to -0.02 per SD increase in cotinine levels), FEV-1 (β = -0.06 litres, 95% CI -0.11 to -0.02 per SD increase in cotinine levels), and FVC (β = -0.05 litres, 95% CI -0.10 to -0.01 per SD increase in cotinine levels). There was also evidence that higher cotinine levels cause increased heart rate (β = 0.75 bpm, 95% CI 0.10 to 1.39 per SD increase in cotinine levels). These results were consistently in the same direction (with the exception of the modal estimation for BMI). There was no clear evidence to suggest a causal effect of cotinine on risk of CHD or COPD.

The results did not substantially differ from the results when the more stringent threshold of $p < 5\times10^{-8}$ was used (Appendices 11 and 12).

Table 5.2. Univariable Mendelian randomisation analysis of cigarettes per day (CPD) and smoking related health outcomes among ever and current smokers (n = 54 SNPs).

Health Outcome	MR method	Ever smokers			Current smokers		
		Effect	95% CI	*p*	Effect	95% CI	*P*
BMI	IVW	-0.09	-0.42, 0.23	0.578	-1.52	-2.15, -0.90	<0.001
	MR-Egger	-0.66	-1.13, -0.19	0.006	-2.14	-3.09, -1.19	<0.001
	W. median	-0.32	-0.57, -0.07	0.013	-2.49	-3.00, -1.98	<0.001
	W. mode	-0.26	-0.46, -0.06	0.010	-2.63	-3.11, -2.15	<0.001
COPD	IVW	1.58	0.96, 2.21	<0.001	2.57	1.64, 3.50	<0.001
	MR-Egger	0.98	0.03, 1.93	0.044	1.78	0.36, 3.21	0.014
	W. median	2.25	1.58, 2.91	<0.001	3.83	2.66, 5.00	<0.001
	W. mode	2.39	1.79, 2.99	<0.001	3.82	2.68, 4.96	<0.001
FEV-1	IVW	-0.16	-0.22, -0.10	<0.001	-0.29	-0.40, -0.17	<0.001
	MR-Egger	-0.09	-0.17, 0.00	0.060	-0.21	-0.39, -0.03	0.020
	W. median	-0.23	-0.28, -0.19	<0.001	-0.44	-0.54, -0.34	<0.001
	W. mode	-0.22	-0.26, -0.18	<0.001	-0.45	-0.54, -0.35	<0.001
FVC	IVW	-0.13	-0.19, -0.06	<0.001	-0.19	-0.31, -0.08	<0.001
	MR-Egger	-0.01	-0.10, 0.08	0.753	-0.06	-0.23, 0.11	0.512
	W. median	-0.15	-0.21, -0.10	<0.001	-0.28	-0.40, -0.16	<0.001
	W. mode	-0.14	-0.19, -0.09	<0.001	-0.27	-0.39, -0.14	<0.001
CHD	IVW	0.11	-0.04, 0.27	0.152	0.11	-0.14, 0.35	0.396
	MR-Egger	-0.16	-0.39, 0.06	0.151	-0.22	-0.58, 0.15	0.250
	W. median	0.16	0.00, 0.32	0.049	0.14	-0.18, 0.45	0.392
	W. mode	0.10	-0.04, 0.24	0.149	0.11	-0.18, 0.39	0.454
Heart Rate	IVW	1.35	0.71, 1.98	<0.001	3.03	1.78, 4.29	<0.001
	MR-Egger	0.56	-0.39, 1.52	0.248	1.91	-0.01, 3.83	0.051
	W. median	1.57	0.99, 2.15	<0.001	4.88	3.56, 6.20	<0.001
	W. mode	1.68	1.11, 2.25	<0.001	4.82	3.57, 6.08	<0.001

Note: A *p*-threshold of 5×10^{-8} was used to determine the single nucleotide polymorphisms (SNPs) associated with smoking heaviness (cigarettes per day [CPD]). BMI = body mass index; COPD = chronic obstructive pulmonary disease; FEV-1 = forced expiratory volume in 1 second; FVC = forced vital capacity; CHD = chronic heart disease; IVW = inverse variance weighted; w. median = weighted median; w. mode = weighted mode. Effects are betas for continuous variables (BMI, FEV-1, FVC and HR) and log odds ratios for binary outcomes (COPD and CHD) per standard deviation increase in number of cigarettes per day.

5.4.2.2 Total effects of smoking heaviness on health outcomes. The total effects of smoking heaviness on health outcomes among ever smokers are displayed in Table 5.2. Among ever smokers, there was no clear evidence of an effect of smoking heaviness on BMI in the IVW analysis, but there was consistent evidence of increased smoking heaviness causing decreased BMI (MR-Egger β = -0.66 kg/m^2, 95% CI -1.13 to -0.19 per SD increase in cigarettes smoked per day) across the other methods. There was also evidence of increased smoking heaviness causing decreased FEV-1 (β = -0.16 litres, 95% CI -0.22 to -0.10 per SD increase in cigarettes smoked per day) and FVC (β = -0.13 litres, 95% CI -0.19 to -0.06 per SD increase in cigarettes smoked per day). There was consistent evidence to suggest that increased smoking heaviness causes increased risk of COPD (lnOR = 1.58, 95% CI 0.96 to 2.21 per SD increase in cigarettes smoked per day). The effect of smoking heaviness on risk of CHD was less clear, with inconsistent evidence across methods (weighted median lnOR = 0.16, 95% CI 0.00 to 0.32 per SD increase in cigarettes smoked per day, but there was no clear evidence of an effect across the other methods). There was consistent evidence to suggest that increased smoking heaviness causes increased heart rate across all methods (β = 1.35 bpm, 95% CI 0.71 to 1.98 per SD increase in cigarettes smoked per day) except for MR-Egger.

The total effects of smoking heaviness on health outcomes among current smokers are displayed in Table 5.2. Among current smokers, there was strong evidence to suggest that increased smoking heaviness causes lower BMI (β = -1.52 kg/m^2, 95% CI -2.15 to -0.90 per SD increase in cigarettes smoked per day), FEV-1 (β = -0.29 litres, 95% CI -0.40 to -0.17 per SD increase in cigarettes smoked per day) and FVC (β = -0.19 litres, 95% CI -0.31 to -0.08 per SD increase in cigarettes smoked per day) and increased heart rate (β = 3.03 bpm, 95% CI 1.78 to 4.29 per SD increase in cigarettes smoked per day) and risk of COPD (lnOR = 2.57, 95% CI 1.64 to 3.50 per SD increase in cigarettes smoked per day). There was no clear evidence of an effect of smoking heaviness on CHD.

5.4.2.3 Tests of directional pleiotropy. MR-Egger intercepts can be used to indicate directional pleiotropy (i.e., evidence that the effect estimate has been biased in a particular direction). Where the intercept value differs from 0, there is evidence of directional pleiotropy (Bowden et al., 2015). There was some evidence of directional pleiotropy in the MR-Egger intercept analysis of cotinine and FEV-1 for ever smokers

(Table 5.3). There was some evidence of directional pleiotropy in the MR-Egger intercept analysis of smoking heaviness and all outcomes except COPD (Table 5.3).

There was no clear evidence of directional pleiotropy in the MR-Egger intercept analysis of cotinine and all health outcomes among current smokers (Table 5.3). There was some evidence of directional pleiotropy in the MR-Egger intercept analysis of smoking heaviness and both FVC and CHD among current smokers (Table 5.3).

The results of the MR-Egger analyses were mostly consistently in the same direction with the other methods. All of the methods provided estimates in the opposite direction to the MR-Egger estimation for the effect of smoking heaviness on CHD among ever and current smokers, however, the confidence intervals crossed the null in all but one estimate.

Table 5.3. MR-Egger test of directional pleiotropy among ever and current smokers.

		Ever smokers			Current smokers		
Exposure	Outcome	Intercept	95% CI	*p*-value	Intercept	95% CI	*p*-value
Cotinine	BMI	0.000	-0.13, 0.13	0.999	0.211	-0.11, 0.53	0.198
	COPD	-0.104	-0.45, 0.25	0.560	-0.081	-0.65, 0.49	0.778
	FEV-1	0.021	0.00, 0.04	0.036	0.025	-0.01, 0.06	0.194
	FVC	0.012	0.00, 0.03	0.118	0.010	-0.01, 0.03	0.337
	CHD	-0.006	-0.06, 0.05	0.835	0.005	-0.12, 0.13	0.928
	Heart Rate	-0.110	-0.35, 0.13	0.370	-0.382	-0.86, 0.09	0.115
CPD	BMI	0.027	0.01, 0.04	0.002	0.029	-0.01, 0.06	0.096
	COPD	0.029	-0.01, 0.06	0.101	0.038	-0.01, 0.09	0.155
	FEV-1	-0.004	-0.01, 0.00	0.029	-0.004	-0.01, 0.00	0.259
	FVC	-0.005	-0.01, 0.00	0.001	-0.006	-0.01, 0.00	0.043
	CHD	0.013	0.01, 0.02	0.001	0.015	0.00, 0.03	0.024
	Heart Rate	0.037	0.00, 0.07	0.035	0.054	-0.02, 0.12	0.133

Note: Cotinine single nucleotide polymorphisms (SNPs) were selected for inclusion based on a threshold of $p < 5\times10^{-6}$. A *p*-threshold of 5×10^{-8} was used to determine the SNPs associated with smoking heaviness (cigarettes per day [CPD]). BMI = body mass index; COPD = chronic obstructive pulmonary disease; FEV-1 = forced expiratory volume in 1 second; FVC = forced vital capacity; CHD = chronic heart disease; 95% CI = 95% confidence interval.

5.4.3 Multivariable Mendelian Randomisation

The results of the MVMR analysis using summary data exploring the direct effects of cotinine and cigarettes per day (n = 54 SNPs) on health outcomes are displayed in Table 5.4 for ever smokers, current smokers, former smokers, and never smokers using the IVW method. Results are presented per SD increase in the exposure phenotype (i.e., cotinine/cigarettes per day).

The Cochran's Q statistics were greater than the number of SNPs included (N = 54) in the majority of the models, indicating heterogeneity. This could be due to horizontal pleiotropy, therefore I repeated the main analysis using the MVMR-Egger method (Table 5.5), a pleiotropy robust method which corrects for measured and unmeasured pleiotropy (Rees, Wood, & Burgess, 2017). However, the statistical power of MVMR-Egger analysis is limited, and low statistical power may be an issue in this analysis.

Table 5.4. Multivariable Mendelian randomisation IVW analysis of cotinine and cigarettes per day (CPD) and smoking related health outcomes among ever, current, former and never smokers (n = 54 SNPs).

	Smoking	Cotinine			CPD			
Outcome	Status	Effect	95% CI	*p*-value	Effect	95% CI	*p*-value	*Q*
BMI	Ever	0.04	-0.17, 0.24	0.725	0.10	-0.47, 0.68	0.721	263.00
	Current	0.05	-0.24, 0.34	0.723	-1.81	-2.64, -0.98	<0.001	118.22
	Former	0.04	-0.16, 0.25	0.658	0.68	0.10, 1.26	0.022	206.09
	Never	-0.04	-0.21, 0.13	0.639	0.96	0.46, 1.45	<0.001	228.42
COPD	Ever	0.03	-0.21, 0.28	0.783	1.99	1.28, 2.70	<0.001	56.28
	Current	0.11	-0.28, 0.51	0.564	3.38	2.27, 4.49	<0.001	42.21
	Former	-0.02	-0.33, 0.30	0.915	1.31	0.40, 2.22	0.006	64.52
	Never	0.25	-0.18, 0.67	0.253	-1.54	-2.75, -0.33	0.013	43.21
FEV-1	Ever	-0.01	-0.03, 0.02	0.659	-0.22	-0.29, -0.15	<0.001	109.57
	Current	-0.02	-0.07, 0.03	0.447	-0.34	-0.48, -0.20	<0.001	84.01
	Former	0.00	-0.03, 0.03	0.951	-0.19	-0.26, -0.11	<0.001	98.14
	Never	0.01	-0.02, 0.03	0.575	-0.03	-0.09, 0.03	0.331	97.75
FVC	Ever	0.00	-0.04, 0.03	0.779	-0.19	-0.28, -0.09	<0.001	130.57
	Current	-0.02	-0.08, 0.04	0.492	-0.24	-0.41, -0.06	0.008	84.25
	Former	0.00	-0.04, 0.03	0.939	-0.17	-0.28, -0.07	0.001	118.94
	Never	0.01	-0.02, 0.04	0.691	-0.05	-0.13, 0.04	0.281	118.86
CHD	Ever	-0.03	-0.13, 0.06	0.459	0.31	0.05, 0.58	0.021	125.08
	Current	-0.02	-0.17, 0.12	0.742	0.30	-0.11, 0.70	0.145	64.77
	Former	-0.04	-0.14, 0.07	0.468	0.31	0.02, 0.61	0.036	119.44
	Never	-0.06	-0.14, 0.02	0.156	0.01	-0.21, 0.24	0.909	64.61
Heart rate	Ever	0.03	-0.30, 0.36	0.857	1.83	0.88, 2.79	<0.001	115.13
	Current	0.50	-0.06, 1.05	0.079	3.00	1.41, 4.59	<0.001	66.34
	Former	-0.15	-0.51, 0.21	0.412	1.42	0.40, 2.44	0.007	107.19
	Never	-0.16	-0.52, 0.20	0.377	-0.06	-1.10, 0.97	0.902	174.15

Note: A *p*-threshold of 5×10^{-8} was used to determine the single nucleotide polymorphisms (SNPs) associated with CPD. A lower threshold of 5×10^{-6} was used to determine the SNPs associated with cotinine due to the low number of SNPs associated at the 5×10^{-8} threshold. BMI = body mass index; COPD = chronic obstructive pulmonary disease; FEV-1 = forced expiratory volume in 1 second; FVC = forced vital capacity; CHD = chronic heart disease. Effects are betas for continuous variables (BMI, FEV-1, FVC and HR) and log odds ratios for binary outcomes (COPD and CHD) per standard deviation increase in cotinine levels/number of cigarettes per day. The Cochran Q statistic tests for instrument strength and validity in the two sample summary data setting.

Table 5.5. Multivariable Mendelian randomisation Egger analysis of cotinine and cigarettes per day (CPD) and smoking related health outcomes among ever, current, former and never smokers (n = 54 SNPs).

	Smoking	Cotinine			CPD		
Outcome	Status	Effect	95% CI	*p*-value	Effect	95% CI	*p*-value
BMI	Ever	-0.01	-0.25, 0.23	0.926	-1.02	-2.01, -0.03	0.044
	Current	0.01	-0.34, 0.36	0.952	-3.38	-4.82, -1.93	<0.001
	Former	-0.01	-0.25, 0.23	0.939	-0.46	-1.46, 0.54	0.368
	Never	-0.11	-0.31, 0.09	0.277	-0.03	-0.89, 0.82	0.941
COPD	Ever	0.01	-0.28, 0.31	0.935	1.35	0.05, 2.65	0.042
	Current	0.00	-0.51, 0.51	0.996	3.80	1.56, 6.03	<0.001
	Former	0.00	-0.38, 0.37	0.996	0.36	-1.29, 2.02	0.668
	Never	0.24	-0.31, 0.78	0.390	-2.88	-5.28, -0.49	0.018
FEV-1	Ever	0.00	-0.03, 0.03	0.888	-0.11	-0.24, 0.01	0.081
	Current	-0.01	-0.07, 0.05	0.684	-0.19	-0.45, 0.07	0.149
	Former	0.01	-0.02, 0.04	0.668	-0.10	-0.23, 0.03	0.145
	Never	0.01	-0.02, 0.03	0.569	-0.01	-0.12, 0.10	0.861
FVC	Ever	0.00	-0.04, 0.04	0.856	-0.02	-0.19, 0.15	0.813
	Current	-0.02	-0.09, 0.05	0.544	0.06	-0.24, 0.37	0.696
	Former	0.01	-0.03, 0.05	0.642	-0.04	-0.22, 0.15	0.704
	Never	0.01	-0.03, 0.04	0.739	-0.03	-0.19, 0.13	0.710
CHD	Ever	-0.07	-0.17, 0.04	0.236	0.02	-0.46, 0.50	0.940
	Current	-0.15	-0.30, 0.01	0.066	-0.30	-1.01, 0.42	0.417
	Former	-0.04	-0.16, 0.08	0.494	0.10	-0.44, 0.64	0.711
	Never	-0.06	-0.16, 0.03	0.175	-0.21	-0.62, 0.20	0.319
Heart	Ever	-0.13	-0.52, 0.26	0.506	0.61	-1.11, 2.33	0.486
rate	Current	0.46	-0.19, 1.12	0.167	0.36	-2.44, 3.15	0.803
	Former	-0.32	-0.74, 0.09	0.130	1.10	-0.79, 2.99	0.253
	Never	-0.46	-0.87, -0.06	0.024	-1.64	-3.47, 0.20	0.081

Note: A *p*-threshold of 5×10^{-8} was used to determine the single nucleotide polymorphisms (SNPs) associated with CPD. A lower threshold of 5×10^{-6} was used to determine the SNPs associated with cotinine due to the low number of SNPs associated at the 5×10^{-8} threshold. BMI = body mass index; COPD = chronic obstructive pulmonary disease; FEV-1 = forced expiratory volume in 1 second; FVC = forced vital capacity; CHD = chronic heart disease. Effects are betas for continuous variables (BMI, FEV-1, FVC and HR) and log odds ratios for binary outcomes (COPD and CHD) per standard deviation increase in cotinine levels/number of cigarettes per day. The Cochran Q statistic tests for instrument strength and validity in the two sample summary data setting.

5.4.3.1 Instrument strength. Instrument strength was calculated using the adapted F-statistic for use with MVMR (Sanderson et al., 2020). The two-sample conditional F-statistic indicates instrument strength of each exposure when accounting for the prediction of other exposures in the model (i.e., whether the SNPs jointly predict smoking heaviness after predicting cotinine levels). As a general rule-of-thumb, F-statistics above 10 indicate strong instruments. This indicated that the SNPs included in the analysis are strong instruments for assessing the direct effects of smoking heaviness while accounting for the effect of cotinine levels (F = 21.66). However, this also indicated that the SNPs may not be strongly associated with cotinine levels while accounting for the effect of smoking heaviness (F = 6.83). Therefore, I cannot be as confident in the estimate of the direct effect of nicotine on smoking-related health outcomes. Nevertheless, I can compare the total and direct effects of smoking heaviness to evaluate whether the total effects are still present when taking nicotine into account; if they differ in magnitude or direction, this gives an indication that nicotine is having a substantial direct effect on the health outcome. These conditional F-statistics were calculated with the use of a less stringent threshold ($p < 5x10^{-6}$) for the inclusion of SNPs associated with cotinine levels; use of the less stringent threshold improved the instrument strength compared with the genome-wide significant threshold by adding more independent SNPs that are only associated with cotinine and not with smoking heaviness ($p < 5x10^{-8}$, F for smoking heaviness = 17.53; F for cotinine = 3.36).

5.4.3.2 Direct effects of cotinine levels on health outcomes. When taking into account the effect of other constituents of cigarette smoke, there was some weak evidence to suggest that increased cotinine levels cause increased heart rate among current smokers (β = 0.50 bpm, 95% CI -0.06 to 1.05 per SD increase in cotinine levels) in the IVW analysis. However, there was no clear evidence of this among ever or former smokers (Table 5.4) indicating some evidence of an acute effect only. There was no clear evidence of any other effect of cotinine levels on smoking-related health outcomes in the IVW analysis (Table 5.4). In the MVMR-Egger analysis, the results were similar but there was weak evidence to suggest that cotinine lowers the risk of CHD among current smokers (Table 5.5) and there was no clear evidence of an effect of cotinine on heart rate except from among never smokers, which is suggestive of pleiotropy (Table 5.5).

5.4.3.3 Direct effects of smoking heaviness on health outcomes. When taking into account the effect of cotinine levels in the IVW analysis, there was no clear evidence to suggest an effect of increased smoking heaviness on BMI among ever smokers (Table 5.4), but there was evidence to suggest that increased smoking heaviness decreases BMI among current smokers (β = -1.81 kg/m^2, 95% CI -2.64 to -0.98 per SD increase in cigarettes smoked per day) and increases BMI among former smokers (β = 0.68 kg/m^2, 0.10 to 1.26 per SD increase in cigarettes smoked per day). Interestingly, I also found that genetic propensity for increased smoking heaviness increases BMI among never smokers which may indicate horizontal pleiotropy. The results of the MVMR-Egger analysis were similar for current smokers, but there was evidence to suggest that that increased smoking heaviness decreases BMI among ever smokers (β = -1.02 kg/m^2, 95% CI -2.01 to -0.03 per SD increase in cigarettes smoked per day), and there was no clear evidence among former and never smokers.

When taking into account the effect of cotinine levels in the IVW analysis, there was evidence to suggest that increased smoking heaviness causes increased risk of COPD among ever smokers (lnOR = 1.99, 95% CI 1.28 to 2.70 per SD increase in cigarettes smoked per day), current smokers (lnOR = 3.38, 95% CI 2.27 to 4.49 per SD increase in cigarettes smoked per day) and former smokers (lnOR = 1.31, 0.40 to 2.22 per SD increase in cigarettes smoked per day). Interestingly, the opposite effect was found among never smokers (lnOR = -1.54, 95% CI -2.75 to -0.33 per SD increase in cigarettes smoked per day) which is indicative of horizontal pleiotropy or collider bias. The results of the MVMR-Egger were similar for ever, current and never smokers, but there was little evidence of an effect among former smokers (Table 5.5).

There was evidence in the IVW analysis to suggest that increased smoking heaviness causes decreased FEV-1 and FVC among ever smokers (β = -0.22 litres, 95% CI -0.29 to -0.15; β = -0.19 litres, 95% CI -0.28 to -0.09 per SD increase in cigarettes smoked per day respectively), current smokers (β = -0.34 litres, 95% CI -0.48, -0.20; β = -0.24 litres, 95% CI -0.41 to -0.06 per SD increase in cigarettes smoked per day respectively) and former smokers (β = -0.19 litres, 95% CI -0.26 to -0.11; β = -0.17 litres, 95% CI -0.28 to -0.07 per SD increase in cigarettes smoked per day respectively). There was no clear evidence to suggest an effect of smoking heaviness on FEV-1 and FVC among never smokers (Table 5.4). However, there was no clear effect of smoking heaviness on FEV-1

or FVC in the MVMR-Egger analysis except for FEV-1 among ever smokers (β = -0.11 litres, 95% CI -0.24 to 0.01 per SD increase in cigarettes smoked per day; Table 5.5).

There was evidence to suggest that increased smoking heaviness causes increased risk of CHD among ever smokers (lnOR = 0.31, 95% CI 0.05 to 0.58 per SD increase in cigarettes smoked per day) and former smokers (lnOR = 0.31, 95% CI 0.02 to 0.61 per SD increase in cigarettes smoked per day), but not among current or never smokers (Table 5.4). There was no clear evidence of an effect of smoking heaviness on CHD among ever, current, former, or never smokers in the MVMR-Egger analysis (Table 5.5).

There was also evidence to suggest that increased smoking heaviness increases heart rate among ever smokers (β = 1.83 bpm, 95% CI 0.88 to 2.79 per SD increase in cigarettes smoked per day), current smokers (β = 3.00 bpm, 95% CI 1.41 to 4.59 per SD increase in cigarettes smoked per day) and former smokers (β = 1.42 bpm, 95% CI 0.40 to 2.44 per SD increase in cigarettes smoked per day), but not never smokers (Table 5.4). In the MVMR-Egger analysis however, there was only weak evidence an increased genetic propensity to heavier smoking decreases heart rate (β = -1.62 bpm, 95% CI -3.47 to 0.20 per SD increase in cigarettes smoked per day).

5.4.4 Comparing the Total and Direct Effects

The total and direct IVW effects of cotinine and smoking heaviness are shown side-by-side for comparison in Figure 5.4 for ever smokers, and Figure 5.5 for current smokers. The results of the MR-Egger and MVMR-Egger results are also shown for comparison in the Appendix 13 for ever smokers and Appendix 14 for current smokers. By comparing the total and direct effects, I am able to evaluate whether the total effects are still present when taking smoking heaviness or nicotine into account; if they differ in magnitude or direction, this gives an indication that the other exposure is having a substantial direct effect on the health outcome.

5.4.4.1 Comparing the total and direct effects on ever smokers. Among ever smokers, the differences between the total and direct effects of smoking heaviness were negligible for the IVW (Figures 5.3) and Egger (Appendix 13) analyses. The effect estimates were similar in magnitude and direction, and the confidence intervals overlapped.

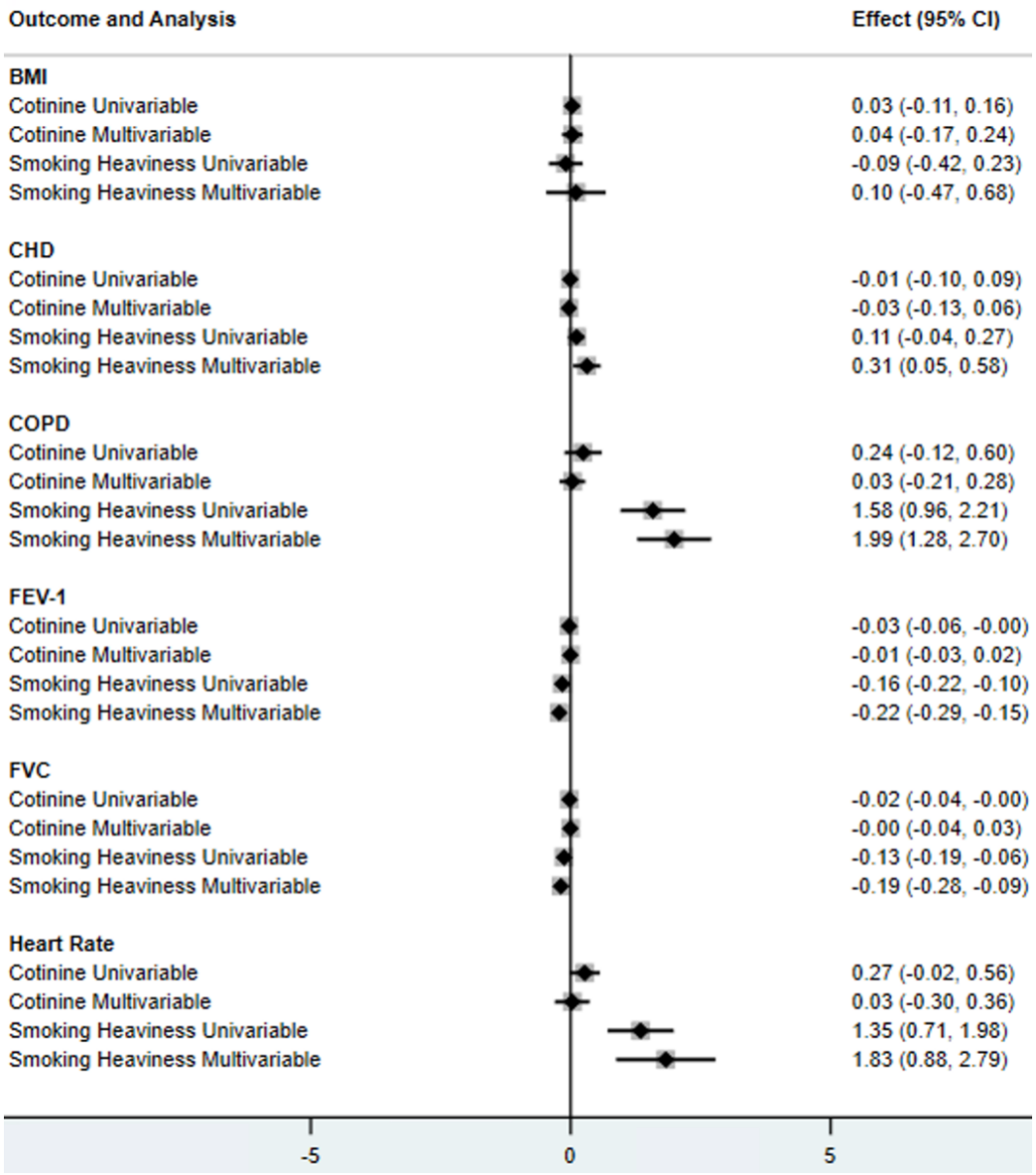


Figure 5.4. Univariable and multivariable Mendelian randomisation analysis of cotinine and smoking heaviness (cigarettes per day) and smoking-related health outcomes among ever smokers (n = 54 SNPs).

Note: A *p*-threshold of 5×10^{-8} was used to determine the single nucleotide polymorphisms (SNPs) associated with CPD. A lower threshold of 5×10^{-6} was used to determine the SNPs associated with cotinine due to the low number of SNPs associated at the 5×10^{-8} threshold. BMI = body mass index; COPD = chronic obstructive pulmonary disease; FEV-1 = forced expiratory volume in 1 second; FVC = forced vital capacity; CHD = chronic heart disease. Effects are betas for continuous variables (BMI, FEV-1, FVC and HR) and log odds ratios for binary outcomes (COPD and CHD) per standard deviation increase in cotinine levels/number of cigarettes per day. Univariable analyses presented are the total effects using the inverse variance weighted (IVW) method. The multivariable analyses reflect the direct effects using the IVW method.

5.4.4.2 Comparing the total and direct effects on current smokers. Among current smokers, the differences between the total and direct effects of smoking heaviness were negligible in the IVW (Figure 5.4) and Egger (Appendix 14) analyses. The effect estimates were similar in magnitude and direction, and the confidence intervals overlapped. Bigger differences in magnitude were seen in the MR-Egger and MVMR-Egger analysis but the confidence intervals still overlapped.

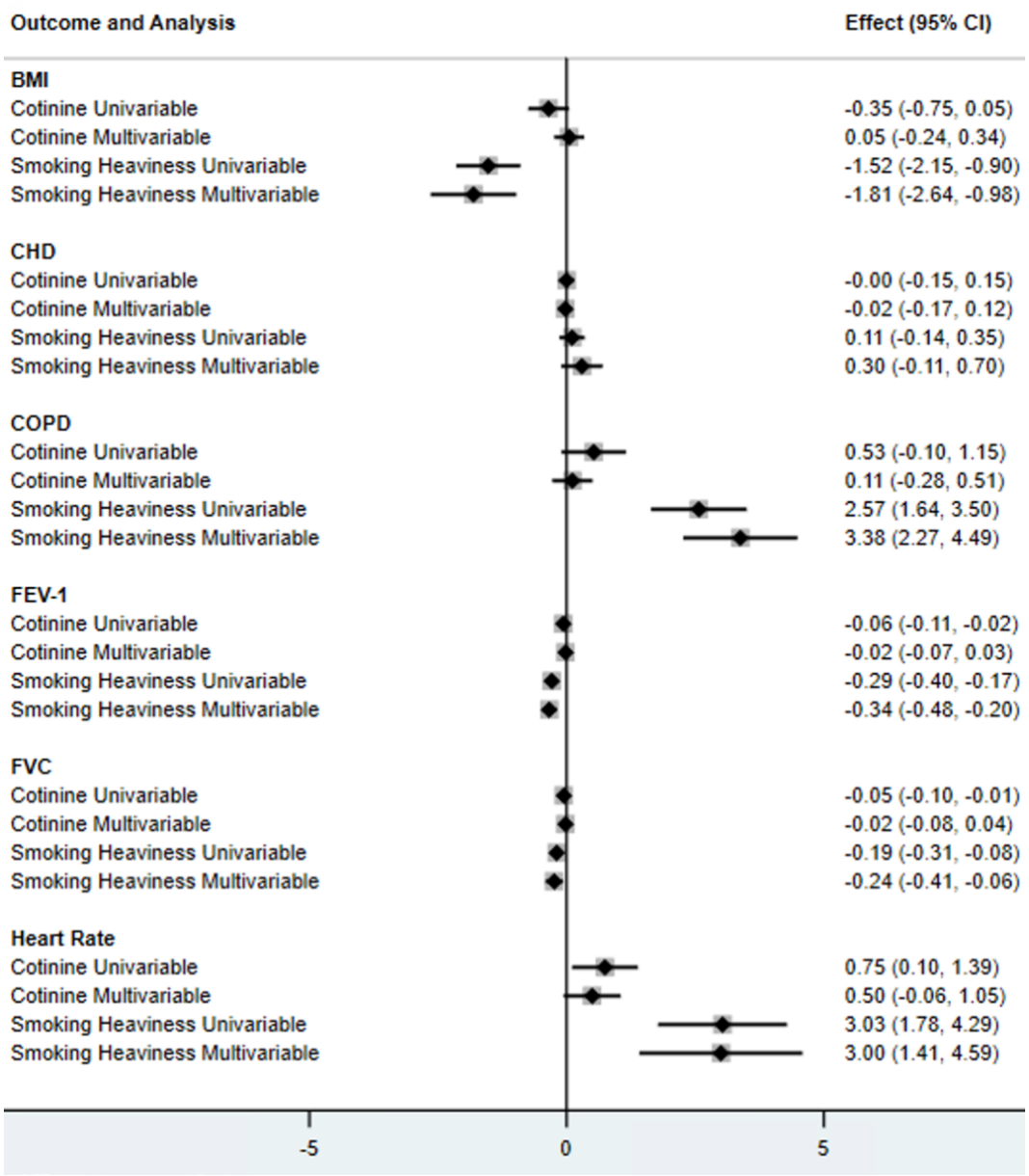


Figure 5.5. Univariable and multivariable Mendelian randomisation analysis of cotinine and smoking heaviness (cigarettes per day) and smoking-related health outcomes among current smokers (n = 54 SNPs).

Note: A *p*-threshold of 5×10^{-8} was used to determine the Single Nucleotide Polymorphisms (SNPs) associated with CPD. A lower threshold of 5×10^{-6} was used to determine the SNPs associated with cotinine due to the low number of SNPs associated at the 5×10^{-8} threshold. BMI = body mass index; COPD = chronic obstructive pulmonary disease; FEV-1 = forced expiratory volume in 1 second; FVC = forced vital capacity; CHD = chronic heart disease. Effects are betas for continuous variables (BMI, FEV-1, FVC and heart rate) and log odds ratios for binary outcomes (COPD and CHD) per standard deviation increase in cotinine levels/number of cigarettes per day. Univariable analyses presented are the total effects using the inverse variance weighted (IVW) method. The multivariable analyses reflect the direct effects using the IVW method.

5.4.5 Sensitivity Analysis

The results did not substantially differ from the results reported when the more stringent threshold of $p < 5\times10^{-8}$ was used (Appendix 15). All the effects described in the IVW MVMR analysis were also evidenced in the sensitivity analyses and, in addition, I found weak evidence of a negative effect of increased genetic propensity for higher cotinine levels on CHD among never smokers (lnOR = -0.13, 95% CI -0.24 to -0.02 per SD increase in cotinine levels).

5.5 Discussion

My findings generally support the known effects of smoking on health. Critically, the direct effects of smoking heaviness are generally similar to the total effect of smoking heaviness, suggesting that these health outcomes are not caused by nicotine (here proxied by cotinine) per se, but by the other non-nicotine constituents of cigarette smoke. In contrast, there is little clear evidence of a direct effect of nicotine on smoking-related health outcomes, although this could be due to a lack of statistical power. Combined, this evidence indicates that nicotine may have little impact on the health outcomes resulting from smoking, whereas the other constituents of cigarette smoke appear to cause numerous health effects related to smoking. This suggests that nicotine use without exposure to cigarette smoke (e.g., via e-cigarettes) could result in fewer negative health consequences related to smoking.

When interpreting these results, it is important to consider the strength of the instruments used. The adjusted F-statistics indicated that the instrument used as a proxy for smoking heaviness was strong, but the instrument used as a proxy for cotinine was weak (Sanderson et al., 2020). In univariable MR, the F-statistic simply indicates the instrument strength of the single exposure; however, the conditional F-statistic in the MVMR context indicates instrument strength of each exposure when accounting for the prediction of other exposures in the model (i.e., whether the SNPs jointly predict smoking heaviness after predicting cotinine levels). Therefore, the conditional F-statistic indicates that I can be confident in the estimate of the direct effect of smoking heaviness when taking into account nicotine exposure (i.e., the effect of other

constituents of tobacco smoke aside from nicotine). However, I cannot be as confident in the estimate of the direct effect of nicotine on smoking-related health outcomes. The main genetic variant identified in the cotinine GWAS (rs10851907) is in LD with rs16969968 (a known functional variant associated with smoking heaviness) which could explain why the instrument is conditionally weak (Ware et al., 2011). For the most part, the differences between the total and direct effects of smoking heaviness were negligible, implying that nicotine has little direct impact on smoking-related health outcomes and that nicotine use without smoke exposure (e.g., via e-cigarettes) is unlikely to result in smoking-related poor health outcomes.

Interestingly, I found evidence of a positive effect of genetic propensity to heavier smoking on BMI among former smokers and those who have never smoked before in the IVW MVMR analysis. Among never smokers, the IVW MVMR results suggest that increased genetic propensity to heavier smoking appears to cause increased BMI, as well as decreased risk of COPD. In the MVMR-Egger results only, there is evidence that increased genetic propensity to heavier smoking appears to decrease heart rate. The effect estimates of these analyses cannot be meaningfully interpreted as never smokers do not smoke any cigarettes per day (despite being predisposed to heavier smoking), but the results (including the high Cochran's Q statistics and weaker evidence of effects in the MVMR-Egger analyses compared with the IVW analyses) are indicative of horizontal pleiotropy (i.e., the genetic variants influencing smoking heaviness also separately influence BMI/COPD through a pathway other than smoking). Furthermore, Taylor and colleagues (2014) also found a positive effect of genetic propensity to heavier smoking on BMI among never smokers. However, stratifying on the exposure (i.e., smoking status) as I have in this analysis could introduce collider bias (Taylor & Munafò, 2014). Predisposition to heavier smoking may influence the likelihood of becoming an established smoker after trying smoking, or the likelihood of an established smoker quitting and becoming a former smoker. Additionally, poor health outcomes may lead smokers to quit smoking; thus, smoking status is a collider, and stratifying by this collider could lead to biased effect estimates indicating a causal effect when there is no true effect. Additionally, selection into UK Biobank could have introduced collider bias; the proportion of smokers in UK Biobank is lower than the UK population, and the mortality rate (indicative of poor health) is also comparatively low. If smoking and poor health result in a lower likelihood of participating in UK Biobank, this would tend to lead to

negatively biased associations whereby we would expect to see a more negative association (i.e., stronger evidence that increased smoking heaviness causes poor health) than would be observed in the intended study population (Munafò et al., 2018). Alternatively, misreporting of smoking status (i.e., reporting to be a never smoker despite being a smoker or former smoker) or second-hand smoke exposure via family members with a shared genetic predisposition to heavier smoking could result in seemingly causal effects among never smokers.

This study is the first to attempt to explore the effects of nicotine use while considering the direct effects of exposure to other constituents of cigarette smoke (and vice versa). I have employed a novel method (MVMR) to explore the causal effect of nicotine on potential health outcomes in order to give an indication of possible future health consequences of long-term nicotine-containing e-cigarette use. However, this study is not without limitations. First, there are issues interpreting findings where the number of cigarettes per day are used as a proxy for smoke exposure. As described by Taylor, Davies, et al. (2014), the number of cigarettes smoked per day is often used to determine lifetime smoke exposure, but there are individual differences in smoking topography (i.e., number of puffs taken per cigarette, average volume per puff etc.) which are not captured by measures of cigarettes smoked per day, meaning measures of cigarettes per day may not adequately capture smoke exposure. Thus, the exclusion restriction assumption is violated as the variants are affecting the outcome via other pathways (although these pathways are still potentially related to smoking) not just the number of cigarettes smoked per day. This may be reflected in the MR-Egger test of directional pleiotropy which indicates that directional pleiotropic effects are observed in the relationship between smoking heaviness and health outcomes, particularly among ever smokers. As there is a smaller sample of current smokers but the size of the intercept is similar to the intercept for ever smokers, there may be some pleiotropic effects for current smokers too but there is insufficient power to detect this as a significant effect. While I have attempted account for horizontal pleiotropy by using pleiotropy robust methods, the statistical power of MVMR-Egger analysis is limited and given that the instrument used as a proxy for cotinine is weak, low statistical power may be an issue in this analysis. Second, cotinine levels are not a perfect measure of nicotine exposure; nicotine metabolism (and therefore cotinine levels) can be affected by a person's age, gender, and even diet (Benowitz et al., 2009) and cotinine could exert

independent physiological effects (Hatsukami et al., 1998; Moran, 2012). Third, BMI (one of the health outcomes of interest) can impact metabolism in general (Wurtz et al., 2014), so could influence nicotine metabolism. Consequently, cotinine levels could differ due to individual differences in metabolism rather than differential exposure to nicotine, so there is likely to be some measurement error in the estimates used to determine which SNPs are independently associated with nicotine consumption. Fourth, the conditional F-statistic indicated that the estimates of the direct effects of nicotine exposure on the health outcomes are likely to suffer from weak instrument bias and should therefore be interpreted with caution. The GWAS of cotinine (Ware et al., 2016) was based on a relatively small sample size (N = 4,548) compared to the GWAS of smoking heaviness (N = 120,744) (Liu et al., 2019) and may have lacked power to detect some influential SNPs.

As more and more GWAS summary data become available, a larger scale GWAS of cotinine may reveal more independent SNPs which can be used as an instrument for nicotine exposure. Future research could extend on these findings using a stronger instrument for cotinine which would allow for clearer interpretation of the causal effect of long-term nicotine use.

5.6 Chapter Summary

Although I found clear evidence of a direct causal effect of exposure to the other constituents of cigarette smoke aside from cotinine on a range of health outcomes, I only observed some evidence of a direct effect of cotinine on increased heart rate and decreased risk of CHD. Despite potential weak instrument bias in the estimates of the direct effect of cotinine, nicotine use via cigarettes appears to have little impact on the selected smoking-related health outcomes because there is little difference between the total effects of smoking heaviness (when cotinine exposure is not taken into account) and the direct effects of smoking heaviness (when the direct effect of cotinine is taken into account). Although nicotine may still have a small influence on health when exposure is independent of smoking, nicotine does not appear to be the primary cause of the negative effects of cigarettes on health. This suggests that long-term use of nicotine without the other constituents of cigarette smoke (e.g., vaping or NRT use) would result in fewer of the selected negative health outcomes than long-term smoking.

However, future studies should further explore the role of nicotine with a stronger instrument (if relevant data become available). In Chapter 6, I further explore the relationship between nicotine and health using experimental methods by investigating the effect of nicotine on food consumption among non-dependent smokers.

Chapter 6 How does acute nicotine exposure affect appetite?

This chapter closely resembles sections from the following pre-registered study protocol:

Ferrar, J., Khouja, J. N., Birch, L., Hamilton-Shield, J., Ness, A., Munafò, M. R., & Attwood, A. S. (2019). Effects of nicotine challenge on eating topography in non-dependent smokers. Retrieved from osf.io/nathz.

I assisted in designing and implementing this experimental study with the support of my co-authors. I took the lead in managing the study and in the analysis and interpretation of the results. Specifically, JF, LB, JHS AN, MM and AA provided feedback on the design of the study.

6.1 Chapter Overview

Although it is widely accepted that there is a causal effect of nicotine on appetite which consequently lowers BMI through reduced food intake, the experimental evidence to support this hypothesis appears to be weak. Evidence suggests that, following nicotine administration, non-smokers display reduced appetite and food intake; however, this effect has not clearly been found among smokers (Perkins et al., 1992; Perkins et al., 1991). In these studies, participants were selected based on their smoking status (non-smokers and dependent smokers). Where samples consist of non-smokers, the effects of nicotine on nicotine-naïve participants can be aversive as nicotine-naïve participants are likely to experience nausea for example (Srivastava, Russell, Feyerabend, Masterson, & Rhodes, 1991), which could result in a reluctance to eat. Where samples consist of dependent smokers, refraining from smoking before an experimental session (in order to control for nicotine exposure) could lead to symptoms of withdrawal which could affect participants' appetite; Jorenby and colleagues (1996) observed increased appetite among participants experiencing nicotine withdrawal following abrupt smoking cessation. Non-dependent smokers should not experience

withdrawal when asked to refrain from smoking before a session and equally should not experience adverse effects of nicotine exposure due to nicotine naivety. In this chapter, I explore the effects of nicotine administration on appetite and subsequent food consumption among non-dependent smokers. This study was disrupted by the outbreak of the COVID-19 pandemic; therefore, the sample size is lower than planned from my power calculation. With this in mind, I refer to this study as a pilot study and present no inferential statistics but outline the planned statistical analyses and present descriptive statistics for the data collected to date. Data collection will resume in the future within the Tobacco and Alcohol Research Group, in order to achieve the pre-planned sample size.

6.2 Introduction

Alongside smoking, obesity is one of the leading risk factors for death across the globe (Ritchie & Roser, 2019), but observational evidence suggests that smoking is associated with having a lower BMI (Audrain-McGovern & Benowitz, 2011). It is widely accepted that nicotine causes weight reduction by reducing appetite (NHS, 2019b). This suggests that using nicotine-containing smoking cessation aids (e.g., e-cigarettes) could prevent post-cessation weight gain (Russo et al., 2016; Russo et al., 2018) which often results in relapse (Mizes et al., 1998). Alternatively, weight gain may be postponed until the individual cuts down their nicotine use via e-cigarettes, but users may be more likely to relapse to vaping rather than relapse to smoking if vaping also reduces their weight. However, as discussed in Chapter 1, the relationship between BMI and smoking appears to be complex; previous evidence from a Mendelian randomisation study suggests that although increased smoking heaviness decreases BMI, increased BMI also appears to causally influence smoking initiation and increases smoking heaviness (Taylor et al., 2019). Combining this with evidence that suggests Mendelian randomisation studies exploring the relationship between BMI and smoking may be affected by horizontal pleiotropy (Taylor, Morris, et al., 2014), demonstrates that further investigation of the relationship is necessary before any clear conclusions can be made about causality.

One hypothesised mechanism by which smoking may influence BMI is through a nicotine-induced reduction in appetite and food consumption. After administering nicotine to non-smokers, Pilhatsch and colleagues (2014) found that subjective ratings

of appetite declined by 17.5%, whereas appetite ratings increased following placebo administration. Similarly, Perkins and colleagues (1991) found participants consumed less food following nicotine administration compared with placebo administration. However, further investigation revealed a clear effect among non-smokers, but not among daily smokers. Another study of male and female daily smokers (using similar methods) found that caloric intake *increased* following nicotine administration compared with placebo administration (Perkins et al., 1992). Although the effect of nicotine on appetite is often presented as fact by trusted sources, such as the NHS (2019b), there is a lack of clear evidence from human experimental studies to support the theory that nicotine reduces appetite and food consumption in smokers.

The mixed findings from experimental studies could be due to lack of statistical power (resulting from small sample sizes) or due to confounding factors influencing the selected samples. For example, nicotine-naïve non-smokers are likely to feel the adverse effects of nicotine (e.g., nausea) as they have not developed a tolerance to nicotine like smokers (Srivastava et al., 1991). This could result in an effect on appetite which is not seen among those who are experienced with nicotine. In contrast, daily smokers develop a tolerance and become dependent on nicotine, thus daily smokers may experience withdrawal symptoms (e.g., increased appetite) when asked to refrain from smoking before a study (in order to standardise nicotine exposure) which could affect their appetite (Jorenby et al., 1996). Non-dependent smokers are experienced with nicotine and therefore should not experience adverse effects when exposed to nicotine and should also not experience any withdrawal symptoms when asked to refrain from smoking. Exploring the effects of nicotine on eating topography among non-dependent smokers could avoid the confounding effects of withdrawal seen among dependent smokers and aversion seen among non-smokers, yet there is currently no published evidence focussed on this subgroup.

The aim of this study was to explore the effect of nicotine on eating topography. I hypothesised that, compared with administration of placebo, administration of nicotine would result in reduced self-selected portion size, amount consumed, and food consumption rate during a subsequent test of food intake.

6.3 Methods

6.3.1 Participants

Participants were non-dependent smokers recruited using posters, web adverts and newsletters. Non-dependent smoking was determined by validated questionnaire (Heatherton, Kozlowski, Frecker, & Fagerstrom, 1991) and was additionally defined as smoking at least five cigarettes per month but not every day. Participants aged between 18 and 50 years of age and who were fluent in English (i.e., English as first language or equivalent level of fluency) were considered eligible for the study. Participants were required to be in good physical and psychiatric health (assessed via self-report).

Those expressing interest in the study were excluded from participation if they were using any other form of nicotine (e.g., e-cigarettes or nicotine replacement therapy) every day or if they had any contraindications for nicotine or placebo use (Appendix 16). Those with self-reported past or present diagnosis of an eating disorder, who were high in dietary restraint according to the Dutch Eating Behaviour Questionnaire (DEBQ) or with a BMI less than 19 kg/m^2 or more than 30 kg/m^2 were excluded from participation. Interested individuals were excluded from participation if they disliked or had allergies/intolerance to the study foods/beverages (pasta, tomato sauce, parmesan cheese, apple juice). Those with self-reported, uncorrected visual or auditory impairment were excluded from participation. Females who were pregnant or breastfeeding were excluded from participation. Pregnancy was assessed by self-report initially and – where participants were uncertain – verified by urine screen on the day of the test session.

On the day of testing, participants experiencing nausea (a rating of 50 or more on a visual analogue scale [VAS] from 0 [not at all] to 100 [extremely]) were asked to reschedule their session. As were participants who failed to adhere to the study requirements: failure to abstain from smoking/nicotine for at least 12 hours (exhaling ≥ 10 parts per million [ppm] carbon monoxide [CO]), failure to abstain from alcohol consumption for at least 24 hours (exhaled breath alcohol concentration > 0 µg/100 ml), or failure to abstain from food or beverages (excluding water and decaffeinated,

unsweetened, coffees and teas without milk) for at least 3 hours prior to the study (self-reported).

After the first 5 participants had completed the study, I made some adjustments to the exclusion criteria and procedure. These adjustments were made in order to maximise recruitment by: (i) offering a wider range of session times, and (ii) not excluding individuals who had abstained from smoking but whose CO readings were between 7 ppm and 10 ppm. The first 5 participants were asked to attend sessions at 12 pm or 1 pm and to abstain from eating anything between breakfast and the test session (approx. 4-5 hours) but due to the change in procedure (additionally offering sessions at 2 pm, 3 pm and 4 pm), the abstinence period was adjusted to 3 hours before the session. Originally, the threshold for CO readings was set to 7 ppm but after a high proportion of participants claimed to have abstained but gave a reading of between 7 ppm and 10 ppm, I decided that the threshold was too sensitive and should be raised to 10 ppm, a commonly used threshold for detecting abstinence from smoking (Ramprasad, Santhosh, & Lim, 2018). The original protocol and amended protocol can both be found on the Open Science Framework (DOI: 10.17605/OSF.IO/XQU8Z). Participants were reimbursed £20 (or equivalent course credit) for completing the study.

6.3.2 Measures and Materials

6.3.2.1 Physical measures. Expired breath alcohol was measured using an AlcoDigital Alcotest 3820 standard device. Expired carbon monoxide was measured using a Bedfont Pico+ Smokerlyser device. Height was measured using a Marsden "Leicester" height measure and weight was measured using an EKS human weighing scale. BMI was calculated as weight (kg)/height $(m)^2$.

6.3.2.2 Questionnaire measures. Dietary restraint was measured using the DEBQ; scores in excess of 2.7 for females or 2 for males indicated high dietary restraint (Domoff, 2015). Nicotine dependence was measured using the Fagerström Test of Nicotine Dependence (FTND) with scores of 3 or more indicating dependence (Heatherton et al., 1991). A 21-item VAS (ranging from 0 [not at all] to 100 [extremely]) was used to determine the extent to which participants were experiencing a range of subjective effects. The subjective ratings related to food and drink consumption (hunger, thirst, fullness, desire to eat, desire to drink, change in taste, excessive salivation),

common nicotine-induced side effects (nausea, dizziness, light-headedness, nervousness, sweatiness, headache, heart racing, confusion, weakness), and feelings of discomfort (lip discomfort, mouth discomfort, throat discomfort, chest discomfort, and stomach discomfort). The Minnesota Nicotine Withdrawal Scale (MNWS) was used to determine the extent to which the participant was experiencing nicotine withdrawal (Hughes & Hatsukami, 1986). MNWS scores range from 0 to 32 and are the sum of eight withdrawal items scored from 0 (none) to 4 (severe).

6.3.2.3 Buccal sprays. We (the research team and I) delivered nicotine via a peppermint flavoured Boots NicAssist 1 mg mouth spray. A dose of 2 mg of nicotine was delivered by administering two sprays of the NicAssist solution. We developed a placebo spray which mimicked the taste, intensity and sensation of the nicotine spray. The spray contained water (88%), Colgate Plax peppermint flavoured mouthwash (3%) and ground black pepper (9%). This nicotine dose and placebo solution were piloted prior to this study on non-smokers (n = 20) and non-dependant smokers (n = 10) to ensure there was some similarity in sensation. The nicotine dose was well tolerated with no adverse events reported. Following the nicotine/placebo spray, we administered a spray of mint breath spray (Smile, Boots UK) in order to mask the unpleasant taste of the nicotine/placebo sprays.

6.3.2.4 Study food and beverages. During the study, participants received a small glass of apple juice (100 ml Copella Cloudy Apple) and a study meal which consisted of 700 g of Sainsbury's basics pasta and 700 g of Sainsbury's tomato and herb Bolognese pasta sauce mixed with 30 g Sainsbury's grated Italian Parmigiano Reggiano D.O.P. cheese. This serving size was selected in order to minimise potential ceiling effects (based on the results of prior food consumption studies conducted at the University of Bristol). The meal was served with 250 ml water and all foods and beverages were stored and prepared in line with the manufacturer's instructions.

6.3.2.5 Eating topography measures. Food portions were weighed on a set of kitchen scales (MyScales, SF-440) and consumption rate was measured using a Mandometer scale (https://mando.se/en/). Consumption rate was additionally calculated manually using the manual weight of food consumed (measured via kitchen scales) and the

in-session recorded meal start and finish time. Satiety (fullness) was measured via the Mandometer mobile phone application.

6.3.3 Design

I employed a repeated measures design with one within-participants factor of nicotine challenge (nicotine vs. placebo). The research assistants and I implemented a single blind procedure of nicotine administration; participants were blinded to the spray they were receiving but the researchers were not blinded due to the unique shaping of the nicotine spray bottle (although branding was obscured from the participants view) and moderate risk of unblinding given the side effects of nicotine administration (e.g., hiccups). The order of administration (i.e., nicotine in the first session and placebo in the second session vs. placebo in the first session and nicotine in the second session) was counter-balanced across participants.

6.3.4 Procedure

Invited participants attended two test sessions at the University of Bristol scheduled at least one week apart. Upon arrival at the laboratory, participants were encouraged to re-read the information sheet (which they had also been sent prior to the day of the test session) before providing their written consent to participate. To ensure participants could make an informed judgement of their eligibility (i.e., to confirm they were not allergic to the study ingredients), a complete list of ingredients and possible side effects for both the nicotine and placebo sprays were provided to the participants. A member of the research team then screened participants to confirm their eligibility for the study. Screening consisted of computerised FTND and DEBQ questionnaire completion, computerised nausea rating, pregnancy test (for females unsure of their current pregnancy status), alcohol breath test, height and weight measures (in order to calculate BMI), CO breath test, and verbal screening of inclusion and exclusion criteria (including adverse events and concomitant medication use). Participants who failed to meet the eligibility requirements were excluded from participating in the study at this point and did not receive any reimbursement. Participants who failed to abstain from smoking, alcohol use or food and drink consumption, or who felt unwell/nauseous on the day of the test session were asked to reschedule.

Eligible participants were randomised to receive either nicotine (2 mg) or placebo in the first session and provided basic demographic information (gender, age). They then completed the MNWS and baseline questionnaires (VAS) on the computer. Following questionnaire completion, the participants received either the nicotine or placebo spray (dependent on their condition allocation). The participants were instructed to hold open their mouth while the researcher delivered three sprays to their buccal cavity (between the participant's cheek and lower gum). The first two sprays were either the nicotine or placebo spray which was immediately followed by one spray of the mint mouth spray. Participants were instructed to close their eyes during the process and were asked to keep their mouth open and to not swallow for a few seconds after the spray. Immediately following the spray administration, participants were given a small glass of apple juice to mask any unpleasant flavours of the sprays. Participants in both conditions were then asked to sit quietly for a 12.5-minute period which allowed the nicotine to be absorbed (in the nicotine condition). We then asked participants to complete the post-spray questionnaires (VAS) on the computer. They then received a large serving bowl of pasta with serving tongs, a bowl, a serviette, a fork and a glass of water. The researcher placed the bowl on the Mandometer in front of the participant and provided the participant with a mobile phone with the Mandometer application open. The researcher gave verbal instructions to the participants informing them to add their desired portion to the bowl and press start on the phone application when they were ready to eat. The participant was told that they could add more food to the bowl at any time during the meal by tapping the 'add more food' option on the Mandometer application, serving themselves their desired portion, and tapping 'start' before continuing to eat. The researcher instructed the participant to tap 'done' in the application when they were finished eating. Participants were prompted to rate their fullness (using a sliding scale from 0% to 100%) every 2 minutes for the duration of their meal via the Mandometer application. The serving bowl containing the study meal was weighed once before the food was served to the participant, and twice after the participant had finished eating to measure: (i) the amount self-served (i.e., including the food the participant had served themselves and the food not consumed), and (ii) measure the amount actually consumed (i.e., excluding any food the participant served themself but did not consume). When participants stated they were finished eating, they completed the post-food computerised questionnaires (VAS). They were then

asked which spray (nicotine or placebo) they believed they had received during their session and how confident they were that they received this spray on a scale of 0 (not confident at all) to 10 (extremely confident) as a manipulation check. The researcher gave the participant a study card with details of the experiment and relevant contact details) to take away with them and reminded them of the date and time of their next session.

On the second study day, the procedure was similar, but the screening procedures were reduced; participants were screened for nausea, pregnancy, failure to abstain from alcohol, smoking, food or drink, adverse events and concomitant medication use. The experimental procedure only differed in that the alternative spray (placebo/nicotine) was delivered, and the participant was debriefed and reimbursed £20 (or equivalent course credit).

6.3.5 Statistical Analysis

6.3.5.1 Sample size calculation. I used the findings of a randomised controlled trial (in which the Mandometer was used to retrain eating behaviour and treat childhood obesity and change in portion size was measured within-groups) to calculate the minimum sample size required for this study (Ford et al., 2009). This study was selected due to the similarity in the outcome measure. Ford and colleagues (2009) found a 45 g (SD = 128 g) mean decrease in self-selected portion size in the treatment group. To achieve 95% power to detect a similar effect size of dz = 0.35, I calculated that I would need to recruit a minimum of 108 participants.

6.3.5.2 Planned analyses. I planned to perform normality checks prior to analysing the data and transform data which were non-normally distributed. I also planned to remove extreme outliers (values which lie more than three times the interquartile range below the lower quartile or above the upper quartile). After completing any necessary transformations and cleaning the data, I intended to explore whether there was a difference between the nicotine and placebo conditions for each of the dependent variables relating to eating topography (self-served portion size, amount consumed, consumption rate) using repeated measures ANOVAs with a within-subjects factor of nicotine challenge (nicotine vs. placebo). I planned to use the same model to analyse the secondary outcome of satiety (change in satiety from baseline). Following the

unadjusted analyses, I intended to re-run the analyses while adjusting for sex, age and BMI. Planned statistical analyses have not yet been performed because the study has been paused due to the COVID-19 pandemic, but descriptive results for the data collected to date are presented. I also intended to explore whether there may be any order effects by re-running the analysis stratified by order of nicotine/placebo administration.

6.3.5.3 Manipulation check. As a manipulation check, I first planned to use chi-square test to explore whether accuracy substantially differed between the placebo and nicotine condition when participants were asked to judge which spray they had received and also compare mean confidence ratings. I then intended to explore whether there were differences over time between the nicotine and placebo group in the VAS data using repeated measures ANOVAs with two within-subjects factors: time (baseline, post-spray, post-food) and nicotine challenge (nicotine, placebo). I expected differences in VAS measures relating to food and drink consumption to be in line with the findings of the main outcomes. I expected VAS ratings relating to common nicotine-induced side effects to be greater following nicotine administration compared with placebo administration. I expected VAS measures relating to discomfort, which may have an impact on eating behaviour, to be similar following placebo and nicotine administration.

6.3.5.4 Planned sensitivity analyses. I planned to conduct sensitivity analyses whereby I repeated the main analysis after excluding individuals who displayed evidence of nicotine withdrawal (i.e., participants scoring 9 or more on MNWS) at baseline. I also intended to run a second sensitivity analysis while excluding individuals who reported an increase of 20 points or more in self-rated nausea between baseline and post-spray VAS questionnaires.

6.4 Results

6.4.1 Disruption due to COVID-19

Data collection commenced on the 13th November 2019 and was due to continue until July 2020. On 20th March 2020, data collection was ceased in response to the COVID-19

pandemic. The study site (University of Bristol) closed on 24th March 2020 and remained closed for human experimental research at the time of my thesis submission. As a result, complete data was collected for only 24 participants. As the sample size was much lower than was required to achieve 90% power (N = 108), I decided that it would be inappropriate to conduct the planned inferential statistics outlined in the methods. Instead, I have analysed the data in line with recommendations for pilot and feasibility studies and only presented descriptive statistics (means [M], standard deviations [SD], and mean differences with 95% confidence intervals where appropriate). Additionally, I was unable to access raw Mandometer data. The Mandometer provides simple graphed data, but raw data is accessed via communication with the Mando data team who were unavailable due to COVID-19.

6.4.2 Participant Characteristics

In total, 28 participants completed session one and 24 participants also completed session two – two participants withdrew after session 1 (indicating a 92% retention rate) and two were unable to complete the second study session due to COVID-19. Of the two participants who withdrew after the first session, both had received nicotine in their first session. The characteristics described reflect the 24 participants who completed session one and two. Six were male (25%) and the average age was 20.25 years of age (ranging from 18 to 31). Nicotine dependence in the sample was low (mean FTND score = 0.17, SD = 0.48), as was nicotine withdrawal at baseline before placebo (mean MNWS score = 5.25, SD = 3.97) and nicotine administration (mean MNWS score = 5.75, SD = 5.27). Scores for restrictive eating were low (M = 1.65, SD = 0.41) and average BMI was 21.88 (SD = 2.49).

6.4.3 Safety

No participants reported adverse events or serious adverse events following drug (nicotine or placebo) administration.

6.4.4 Manipulation

In both the placebo and nicotine condition, 75% of participants accurately guessed which spray had been administered which could suggest the participants were not

adequately blinded. However, the mean rating of confidence in their guess (on a scale of 0 [not at all confident] to 10 [extremely confident]) was only 5.5, which suggests participants were unsure of their condition and therefore were adequately blinded.

6.4.5 Self-Selected Portion Size

Self-selected portion sizes (i.e., the amount of food participants transferred from the serving bowl onto their plate irrespective of the amount of food they actually consumed) in the placebo condition and the nicotine condition are displayed for each participant in Figure 6.1. Following placebo administration, participants served themselves 403 g (SD = 162 g) of pasta on average. Following nicotine administration, participants served themselves 427 g (SD = 141 g) of pasta on average. This corresponds to a mean difference of 24 g (95% CI -29 g to 76 g).

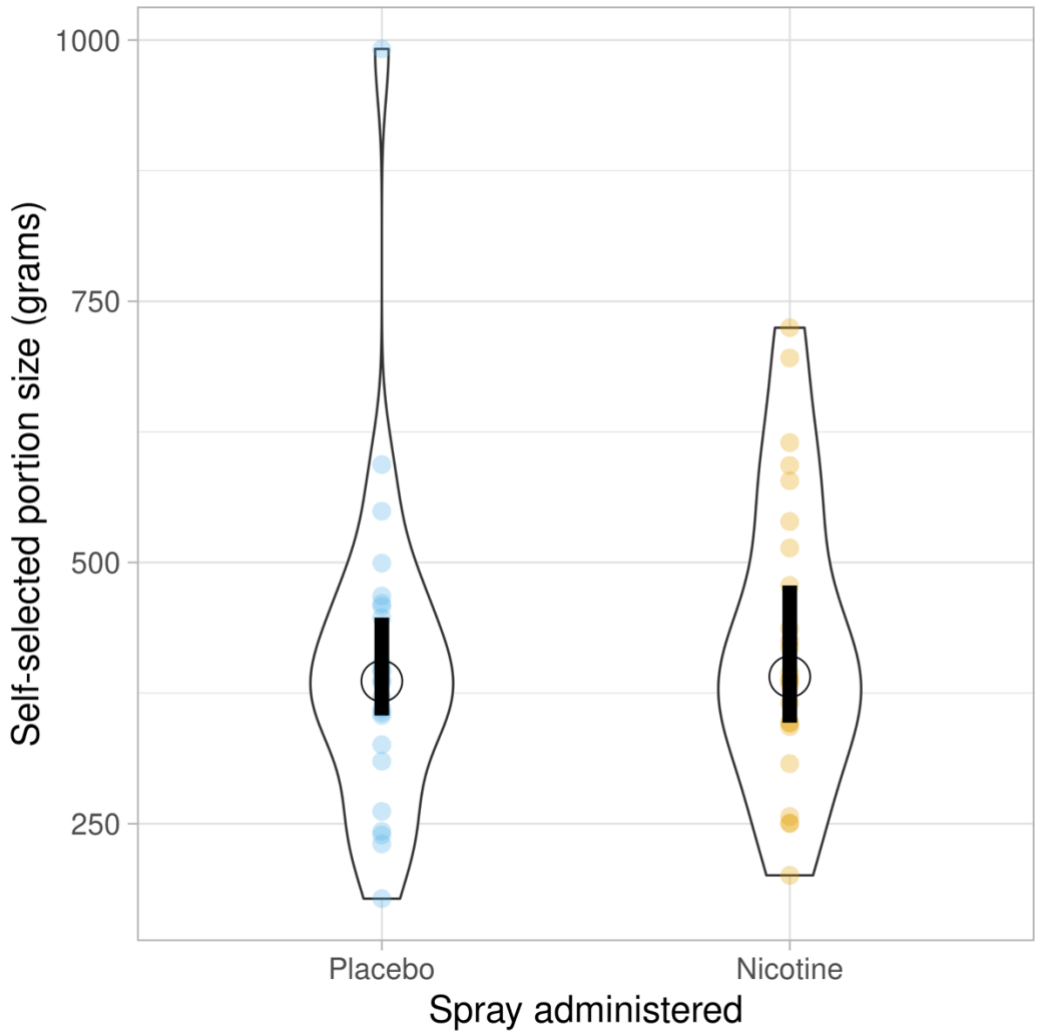


Figure 6.1. Average self-selected food portion (grams) after spray administration.

Note: The summary of the data is shown as a violin plot which reflects individual participant data (solid coloured dots) the data distribution (outer line), and the median (open circle). A vertical black bar indicates the 95% confidence interval determined by bootstrapping.

6.4.6 Food Consumed

The portion of food consumed in grams by each participant in the placebo condition and the nicotine condition are displayed in Figure 6.2. Following placebo administration, participants consumed 399 g (SD = 166 g) of pasta on average. Following nicotine administration, participants consumed 416 g (SD = 141 g) of pasta on average. This corresponds to a mean difference of 18 g (95% CI -33 g to 69 g).

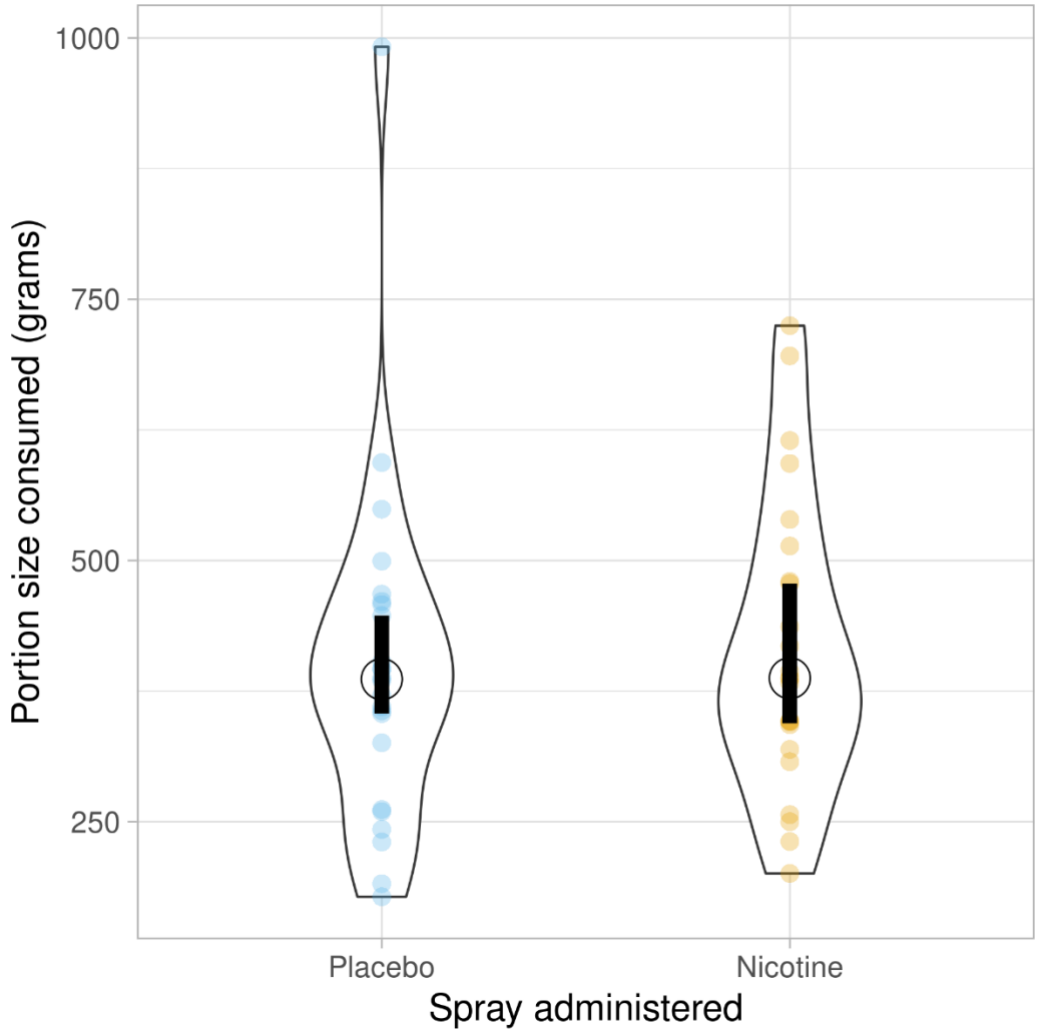


Figure 6.2. Average food portion (grams) consumed after spray administration.

Note: The summary of the data is shown as a violin plot which reflects individual participant data (solid coloured dots) the data distribution (outer line), and the median (open circle). A vertical black bar indicates the 95% confidence interval determined by bootstrapping.

6.4.7 Food Consumption Rate

Consumption rate was measured manually; Mandometer data was unavailable. The food consumption rate for each participant in the placebo condition and the nicotine condition are displayed in Figure 6.3. Following placebo administration, participants consumed 53 g (SD = 17 g) of pasta per minute on average. Following nicotine administration, participants consumed 51 g (SD = 16 g) of pasta per minute on average. This corresponds to a mean difference of -2 g of pasta per minute (95% CI -7 g to 2 g).

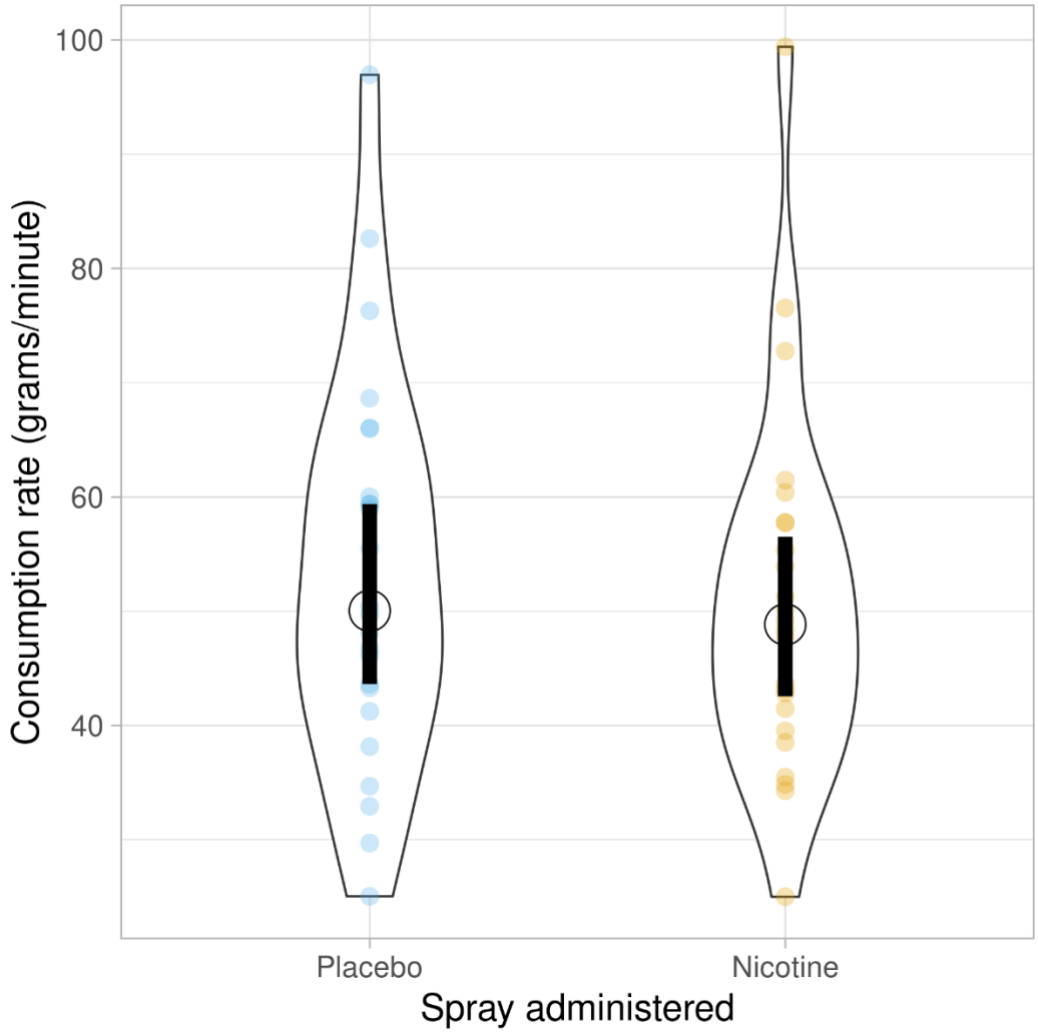


Figure 6.3. Average food consumption rate (grams/minute) after spray administration.

Note: The summary of the data is shown as a violin plot which reflects individual participant data (solid coloured dots) the data distribution (outer line), and the median (open circle). A vertical black bar indicates the 95% confidence interval determined by bootstrapping.

6.4.8 Satiety

There was considerable missing satiety data where participants had missed or ignored prompts to enter their satiety rating. At the start of the meal in the placebo condition, 19 participants rated their satiety, and 21 participants rated their satiety within 2 minutes of finishing their meal. At the start of the meal in the nicotine condition, 20 participants rated their satiety, and 22 participants rated their satiety within 2 minutes of finishing their meal. Only 13 participants had sufficient data to calculate change in satiety following both placebo and nicotine administration. Change in satiety ratings (from meal start time to within 2 minutes of meal completion) in the placebo condition and the nicotine condition are displayed for individuals with complete satiety data (n = 13) in Figure 6.4. Satiety ratings increased by 48% (SD = 16%) on average following food consumption in the placebo condition. Satiety ratings increased by 42% (SD = 6%) on average following food consumption in the nicotine condition. This corresponds to a mean difference of 7% (95% CI -11% to 25%). Satiety ratings increased (i.e., participants became fuller after eating) for all participants following meal consumption, bar one participant whose satiety ratings decreased during meal consumption following nicotine administration (i.e., the participant became less full after eating).

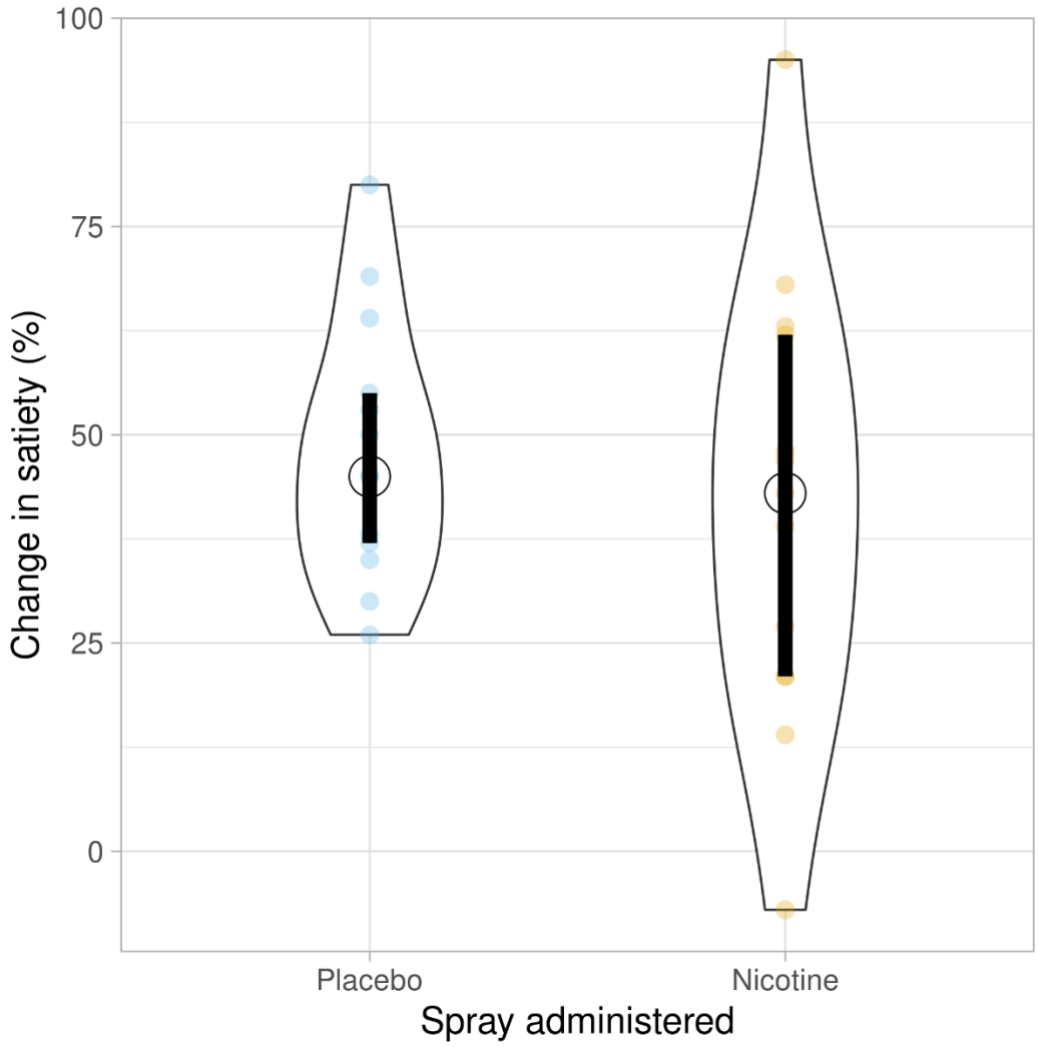


Figure 6.4. Average change in satiety rate (from meal start to within 2 minutes of meal completion) following spray administration.

Note: The summary of the data is shown as a violin plot which reflects individual participant data (solid coloured dots) the data distribution (outer line), and the median (open circle). A vertical black bar indicates the 95% confidence interval determined by bootstrapping.

6.4.9 Subjective Measures

Due to a computing error, baseline subjective data for session two (nicotine condition) was unavailable for one participant. The data reported for baseline subjective measures in the nicotine administration condition relate to the 23 participants with available data. The mean ratings (ranging from 0 [not at all] to 100 [extremely]) at baseline, post-spray (after placebo or nicotine administration) and post-food (after food consumption) in the placebo condition and the nicotine condition are displayed for each subjective measure in Table 1.

Table 6.1. Mean VAS ratings of subjective effects over time (baseline, post-spray, post-food).

	Placebo			Nicotine		
Subjective measure	Baseline	Post-spray	Post-food	Baseline	Post-spray	Post-food
Hunger	66.33 (3.35)	67.38 (4.00)	15.63 (3.93)	65.91 (3.72)	55.63 (4.41)	11.04 (3.99)
Thirst	61.79 (3.27)	56.92 (4.03)	25.17 (3.68)	60.78 (3.29)	49.17 (4.97)	25.63 (3.45)
Fullness	22.00 (3.92)	20.33 (3.80)	80.04 (2.54)	20.87 (4.25)	25.08 (4.16)	77.25 (3.72)
Desire to eat	71.33 (2.88)	68.63 (3.81)	13.38 (2.68)	67.30 (4.23)	61.42 (4.03)	11.88 (3.97)
Desire to drink	69.33 (2.84)	61.50 (3.64)	29.50 (3.54)	72.57 (3.09)	56.08 (4.53)	35.46 (4.85)
Nausea	16.29 (4.24)	13.33 (3.82)	9.96 (3.64)	15.00 (4.13)	22.71 (4.74)	18.42 (4.75)
Dizziness	9.92 (3.12)	7.67 (2.44)	5.29 (2.10)	8.00 (2.43)	22.75 (4.54)	12.25 (3.98)
Light-headedness	15.75 (4.22)	15.46 (4.13)	7.08 (2.69)	13.48 (3.87)	27.46 (4.87)	15.96 (4.36)
Nervousness	16.67 (3.86)	11.29 (3.88)	5.46 (2.68)	12.30 (3.16)	8.50 (2.62)	8.88 (3.14)
Sweatiness	19.29 (4.22)	12.13 (3.05)	7.29 (2.46)	21.30 (4.62)	13.71 (3.80)	10.54 (3.65)
Headache	14.50 (4.72)	13.42 (4.75)	5.50 (1.70)	13.35 (4.52)	20.08 (5.17)	14.54 (4.72)
Salivation	22.75 (4.36)	22.83 (4.51)	16.71 (5.28)	19.91 (4.93)	21.54 (4.75)	15.33 (5.16)
Heart racing	4.54 (1.65)	5.33 (1.91)	4.46 (1.97)	7.26 (2.21)	11.63 (3.66)	9.96 (3.45)
Confusion	5.54 (1.79)	4.71 (1.78)	3.50 (1.98)	7.22 (2.96)	6.75 (2.63)	6.38 (2.03)
Weakness	14.92 (3.65)	11.71 (3.75)	7.21 (2.48)	15.65 (4.42)	20.54 (4.81)	14.46 (4.57)
Taste	8.17 (2.76)	21.71 (4.48)	18.25 (5.19)	10.35 (3.28)	26.29 (5.17)	23.58 (6.11)
Lip discomfort	11.79 (3.57)	10.63 (2.96)	6.33 (1.83)	8.57 (3.57)	14.71 (4.88)	6.88 (2.52)
Mouth discomfort	7.67 (2.54)	15.17 (3.86)	7.63 (2.13)	7.87 (2.73)	23.67 (4.67)	11.50 (3.91)
Throat discomfort	10.83 (3.29)	12.83 (3.73)	8.58 (3.07)	9.57 (3.50)	13.13 (4.77)	7.54 (3.50)
Chest discomfort	5.13 (2.09)	3.71 (1.03)	7.00 (2.51)	9.83 (3.99)	7.33 (2.97)	4.38 (1.78)
Stomach discomfort	14.96 (4.35)	14.29 (4.24)	19.42 (5.29)	14.78 (3.90)	19.67 (5.02)	22.13 (5.24)

6.4.10 Sensitivity Measures

In the proposed sensitivity analyses, I planned to restrict the analysis by MNWS ratings and change in nausea. Five participants (21%) experienced substantial withdrawal (MNWS > 9) in the placebo condition and six (25%) experienced substantial withdrawal in the nicotine condition. Four participants (17%) experienced substantial changes in nausea rating pre- to post-spray in the nicotine condition (increase in ratings of 20 points or more) and no participants experienced substantial changes pre- to post-spray in the placebo condition.

6.5 Discussion

Over a three-month recruitment period, 24 participants completed the pilot study. The study retention rate was high (92%) when excluding those who could not complete the study due to COVID-19 restrictions. The amended study procedures increased the rate of enrolment (5 participants completed the study within six weeks prior to the changes, and 19 participants completed the study within seven weeks following the changes). The quality of data collected for the main outcome was good, however, the data quality for the secondary outcome (satiety) was moderate, and access to data collected via the Mandometer was limited due to COVID-19 restrictions. The data should be sufficient to support my planned main analysis once the pre-determined sample size has been achieved, and with some minor amendments (i.e., introducing a pen and paper or computerised rating of satiety) the data should also be sufficient to support my secondary analysis.

Due to the small sample size and limited statistical power of this pilot study, the results described in this Chapter cannot be used to infer a statistical difference between the placebo and nicotine conditions. Although I discuss the pattern of results in relation to my hypotheses here, it is important to highlight that the specific estimates (mean differences and confidence intervals) reported could change in size and direction when the sample size increases (Button et al., 2013). The observed direction of the effects of nicotine administration on self-selected portion size and food consumption are in contrast to my hypotheses; participants actually served themselves slightly *larger* portions and consumed slightly *more* on average following nicotine administration

compared with placebo administration. The observed direction of the effect of nicotine administration on satiety was also not in line with my hypothesis; participants reported feeling less satiated on average over the course of the study meal following nicotine administration compared with placebo administration. The pattern of these results supports previous evidence suggesting nicotine increases caloric intake among smokers (Perkins et al., 1992). However, the observed direction of effect of nicotine administration on consumption rate was in line with my hypothesis; consumption rate decreased following nicotine administration, with participants consuming slightly less food per minute compared with the placebo administration. Additionally, the pattern of results for the subjective ratings are in line with a nicotine-induced reduction in appetite; mean subjective ratings of hunger decreased on average following nicotine administration whereas mean ratings of hunger increased following placebo administration. The pattern of these results supports previous evidence suggesting nicotine decreases caloric intake among smokers (Perkins et al., 1991).

Although the patterns reported warrant further investigation, I cannot draw any clear conclusions from the current findings without conducting inferential statistics which would be inappropriate with such a small sample size (Button et al., 2013). With a sample size of 24, I only have 38% power to detect an effect of $dz = 0.35$. Consequently, it is unlikely that small effects of nicotine on self-selected portion size, food consumption, or consumption rate would be detected as a result of conducting my planned inferential statistics. When testing can safely resume, I plan to complete testing with a further 88 participants and conduct the full planned analysis with sufficient power. If I find that the results are in line with my hypotheses when the full sample size has been achieved, I will conclude that nicotine appears to reduce appetite among non-dependent smokers and this could explain the observed relationship between smoking and lower BMI (Audrain-McGovern & Benowitz, 2011). If the data do not support my hypotheses, I will conclude that this relationship may not be due to a nicotine-induced reduction in appetite and food consumption.

A further limitation of the study is that the satiety measure was limited by substantial missing data due to participants not completing the questions when prompted. When the study recommences, this will be addressed by introducing clearer instructions regarding the completion of these questions as well as a pen and paper or computerised

rating which participants will complete directly before and after meal consumption. The study was also limited by the inaccessibility of data from the Mandometer due to COVID-19 restrictions. When these data are accessible, I will contact the data team and arrange for the raw data to be sent to the research team at regular intervals so that this issue can be avoided in the event of another disruption.

Even with the current limitations of the study, some strengths of the study are still evident. The patterns in subjective ratings suggest that the dose of nicotine administration was adequate to result in a biological response, with higher ratings of nicotine-related side effects on average in the nicotine compared with the placebo condition – particularly for nausea, dizziness and light-headedness which are common side effects of nicotine administration (as stated on the Boots NicAssist information leaflet). In both the placebo and nicotine condition, 75% of participants accurately guessed which spray had been administered, but the mean rating of confidence in their guess (on a scale of 0 [not at all confident] to 10 [extremely confident]) was only 5.5. Therefore, the participants were adequately blinded given the uncertainty of their judgement. Average ratings of discomfort which could have impacted food consumption appeared to be low in both conditions. However, there may be some differences in mouth discomfort which may have an impact on eating behaviour – this should be explored when sufficient data is available. As participants reported more mouth discomfort but also consumed more food following nicotine administration compared with placebo administration, it does not appear that mouth discomfort induced by the nicotine spray reduced food intake.

6.6 Chapter Summary

In summary, the results show a pattern of reduced consumption rate but increased self-selected portion size and food consumption following nicotine administration compared with placebo administration. Therefore, the relationship between nicotine and BMI may not be as straightforward as previously assumed; the relationship may be a result of factors aside from, or in addition to, nicotine-induced reductions in food consumption. However, no clear conclusions can be drawn until inferential statistics are conducted with a sufficient sample size.

Chapter 7 Discussion

7.1 Summary

The aims of this thesis were three-fold, namely to:

1. Identify which young adults are more likely to use e-cigarettes by describing the demographics and profiles of young e-cigarette users in a UK sample.
2. Explore why young adults use e-cigarettes, including the subjective reasons given by vapers, and the genetic factors influencing vaping.
3. Investigate the potential consequences of e-cigarette use on smoking behaviour and health outcomes.

I highlighted the need to address these aims in Chapter 2, showing that young people who have never smoked before are being exposed to vaping which could lead to negative consequences such as smoking initiation. I addressed my first aim in Chapter 3 and found substantial differences in demographic, behavioural and lifestyle factors between young adults who had ever and never vaped, as well as differences between those who were former and current vapers. I addressed my second aim in Chapters 3 and 4, identifying reasons for using e-cigarettes among a sample of UK young adults who had both ever smoked and ever vaped, and exploring whether there is a shared genetic predisposition to smoking and e-cigarette use. I addressed my third aim in Chapters 2, 3, 5 and 6, investigating whether e-cigarette use could lead to smoking initiation, continued smoking, continued vaping and smoking-related health outcomes. In this Chapter, I discuss the findings of each of my thesis chapters in relation to my aims, the conclusions and implications of the thesis, as well as the strengths, limitations and potential future directions of the work.

7.1.1 Which Young Adults Are More Likely to Use E-cigarettes?

In Chapter 2, I explored the association between e-cigarette use among non-smokers and later smoking in a systematic review and meta-analysis. While reviewing studies which explored this association, I identified a substantial number of non-smoking e-cigarette users. As e-cigarettes are only a form of harm reduction to smokers (Kimber, Frings, Cox, Albery, & Dawkins, 2020), this highlighted the need to identify which young

adults are most likely to vape. In Chapter 3, I identified multiple factors which were associated with ever using e-cigarettes in a cohort of young adults in the UK: young adults who are male, of lower socioeconomic position at birth, who have poorer mental health, who engage in other risky behaviour, and who smoke are more likely to have ever vaped. Current vapers were more likely to have lower parental socioeconomic position at birth, report anxiety and have smoked by the age of 23 years but were less likely to be hazardous or harmful alcohol users than former vapers.

7.1.2 Why Do Young Adults Use E-cigarettes?

After identifying which individuals are more likely to vape, it is important to understand *why* they vape. In Chapter 3, I found that the most common reason given for vaping by 23 years old was out of curiosity (51%), the second most common reason was to quit smoking (32%) and third was vaping to cut down their smoking (23%). This is in contrast to older adults in the UK who are most likely to use an e-cigarette to quit smoking (Action on Smoking and Health, 2019a).

In addition to the subjective reasons given for vaping, my results indicate that vapers may be genetically predisposed to smoking as well as e-cigarette use. In Chapter 4, I found evidence to suggest that there is a potential shared genetic predisposition to smoking initiation and e-cigarette use; I observed associations of a similar magnitude between smoking initiation polygenic risk scores and both self-reported smoking initiation and ever e-cigarette use. Given the strong correlation between the two behaviours, this is not necessarily surprising, and the association found could be due to smoking directly impacting e-cigarette use – a logical assumption given that 32% of young adults in this cohort stated that they have used e-cigarettes to quit smoking (as found in Chapter 3) and 46% used e-cigarettes for any smoking-related reason.

However, I also observed associations with smoking initiation polygenic risk scores and risk-taking and impulsivity, even when using the genome-wide significant *p*-value threshold for the inclusion of genetic variants within the polygenic risk scores. This indicates that there could be horizontal pleiotropic effects (i.e., the genetic variants in the polygenic risk scores may impact e-cigarette use via a route that is not through

smoking). Furthermore, I observed similar associations among never smokers. Rather than specifically predicting smoking or nicotine use, the smoking initiation polygenic risk scores may be capturing a broader risk-taking phenotype. If young adults are genetically predisposed to risk-taking (including smoking and e-cigarette use) this could explain why e-cigarette use *appears* to act as a gateway to smoking among young people (as found in Chapter 2).

7.1.3 What Are the Consequences of Using E-cigarettes?

In Chapter 2, I provided an update to a previous meta-analysis (Soneji, Barrington-Trimis, Wills, Leventhal, et al., 2017) by identifying 17 studies (11 studies in addition to those found by Soneji and colleagues) which explored the association between e-cigarette use among non-smokers and later smoking. By meta-analysing the results of these studies, I found a strong association between e-cigarette use and later smoking, whereby non-smokers who had vaped were 3 times more likely to subsequently ever smoke. Subgrouping and stratification of the results (e.g., stratification by risk of bias) revealed some slight differences between groups, but consistently indicated a strong positive association.

The results could suggest that e-cigarettes act as a gateway to smoking; however, I also identified some clear limitations of the included studies. First, there was a lack of biochemical verification of smoking or vaping status, meaning the estimates could suffer from measurement error and reverse causality. Second, some of the included studies lacked appropriate and adequate adjustment for potential confounding beyond basic demographic factors. Third, nicotine content was only accounted for in one study, despite often being implicated as the mechanism by which e-cigarettes act as a gateway to smoking (Bell & Keane, 2014). Although the majority of studies met three out of four pre-selected Bradford-Hill causality criteria, the criteria I used were relatively relaxed, and other (non-causal) explanations for the association between e-cigarette use and later smoking (e.g., a common liability) cannot be ruled out. Reviewing these studies highlighted a substantial number of non-smoking e-cigarette users who are potentially at greater risk of smoking and other poor health outcomes.

In Chapter 3, different reasons for vaping were associated with the likelihood of continued smoking and vaping among ever smokers and ever vapers. I found that vaping

out of curiosity was associated with lower likelihood of vaping at 24 years and was associated with continued smoking at 24 years among those who smoked just prior to vaping. However, vaping out of curiosity was also associated with an increased likelihood of using neither product compared with being a current smoker at 24 years among the full sample, and there was no clear association between vaping out of curiosity and current smoking at 24 years among individuals who did not regularly smoke just prior to vaping. Therefore, vaping out of curiosity does not appear to encourage continued vaping or smoking uptake among non-current smokers, but does not support smoking cessation either. Vaping to quit smoking at 23 years was associated with higher likelihood of vaping at 24 years and lower likelihood of smoking at 24 years (among those who were smoking just prior to vaping) compared to those not vaping to quit. Similarly, vaping to cut down smoking was associated with continued vaping at 24 years, but in contrast, was associated with continued smoking at 24 years (among those who smoked just prior to vaping). Therefore, I concluded that e-cigarette use alone may not lead to discontinued smoking unless e-cigarettes are used with the intention to quit smoking.

In Chapter 5, I found little evidence of a direct effect of cotinine on smoking-related health outcomes. This implies that the other constituents of tobacco smoke (aside from nicotine) are the cause of the selected smoking-related poor health outcomes. Although there may have been some weak instrument bias affecting the estimates of the direct effect of cotinine, this suggests that nicotine use (e.g., via e-cigarettes) may not lead to poor health outcomes associated with smoking.

I further explored the relationship between nicotine exposure and health in Chapter 6 but did not present any inferential statistics. Instead, I outlined the planned analyses for an experimental exploration of the effects of nicotine administration on eating topography which was disrupted by the COVID-19 pandemic. I presented descriptive statistics for the data collected to date. Although I cannot draw any conclusions at this stage (and the magnitude and direction of the estimates may be subject to change when the study is complete), the pattern of results are in the same direction as only one out of four of my hypotheses. In line with my hypotheses, participants consumed their meals slower following nicotine administration compared with placebo administration. In contrast to my hypotheses, participants served themselves more, consumed more, and

were less full after meal consumption following nicotine administration compared with placebo administration.

7.2 Conclusions and Implications

The results provide further evidence to contribute to the debate on whether e-cigarettes act as a gateway to smoking or could be a result of a shared liability. The results also provide novel evidence regarding the potential health implications of using nicotine without exposure to tobacco smoke (e.g., via e-cigarettes).

7.2.1 The Gateway Effect and Common Liability Theory

In Chapter 2, I found evidence of an association between e-cigarette use and later smoking which could be used to infer that e-cigarettes act as a gateway to smoking. However, I also highlighted the limitations present in the current literature. Consequently, I have refrained from concluding this association is due to a gateway effect. Furthermore, I found evidence to suggest that the strong association found between e-cigarette use and later smoking could be due to a shared liability (Chapters 3 and 4). E-cigarette users share many demographic, behavioural and lifestyle characteristics with smokers (Chapter 3), and also share a genetic predisposition which may reflect a general predisposition to risk-taking (Chapter 4).

There is a strong correlation between risk-taking and curiosity (Sagone & Caroli, 2013), and I found that curiosity was the most common reason for vaping among young adults in the UK who have both vaped and smoked at least once (Chapter 3). Nicotine addiction is a commonly cited explanation for why e-cigarettes may act as a gateway to smoking, yet the young adults who vaped out of curiosity did not appear to develop an addiction to nicotine (i.e., they did not seem to transition from non-current smoking/vaping to current smoking/vaping after trying e-cigarettes). These results could highlight a substantial group of 'experimenters' who try vaping and smoking but do not transition to current use. Given the majority of studies included in my meta-analysis (Chapter 2) explored the association between *ever* vaping and *ever* smoking, the results are likely to capture 'experimenters' who are curious about both vaping and smoking and are likely to take risks. Including these individuals in studies exploring the gateway hypothesis (particularly studies using ever vaping and ever smoking as the exposure and outcome)

may inflate estimates of a gateway effect mediated by nicotine addiction as these 'experimenters' do not appear to be addicted to nicotine.

Ideally, I would have additionally explored the association between reasons for vaping and later vaping and smoking among those who had never smoked at 23 years, however, the sample was not appropriate to explore this; e-cigarettes only became available late in adolescence for this cohort whereas cigarettes were widely available throughout their adolescence, so the young adults had limited opportunity to experiment with e-cigarettes prior to smoking. As a result, I had insufficient power to stratify the data by never smokers at 23 years. Therefore, it is inappropriate to make direct conclusions regarding the gateway hypothesis based solely on the results of Chapter 3. However, triangulating the results and conclusions of Chapters 2, 3 and 4, shows some support for a shared liability between e-cigarette use and smoking.

Of course, the two theories are not mutually exclusive, and it is possible that there is both a gateway effect and a shared liability influencing young adult e-cigarette use. I should also acknowledge that my own views and beliefs may have influenced my interpretation of these results (as I assumed the views and beliefs of other researchers may have impacted their conclusions regarding the gateway hypothesis in Chapter 2).

7.2.2 Potential Health Consequences of E-cigarette Use

In Chapter 2, I found a substantial number (10%) of young people who are non-smokers have ever vaped. To protect these young people from harm, it is vital to understand the health implications of nicotine use without exposure to tobacco smoke. In Chapter 5, I found no clear evidence to suggest that nicotine exposure without exposure to cigarette smoke results in smoking-related health outcomes. This suggests that nicotine use via vaping should not substantially expose non-smoking vapers to smoking-related health risks such as chronic obstructive pulmonary disease.

Although I found limited evidence to suggest that long-term nicotine use results in smoking-related ill health (Chapter 5), this evidence is by no means sufficient to determine that e-cigarettes are safe for non-smokers. The instruments used as a proxy for cotinine levels were weak and nicotine is just one possible constituent of e-cigarette vapour; there are many different ingredients that can be included in e-liquids which can

result in different and potentially harmful emissions (Behar et al., 2016). E-cigarettes could also pose risks which are not associated with smoking, such as EVALI (e-cigarette and vaping product use associated lung injury). Nevertheless, smokers expose themselves to many more chemicals and toxicants in tobacco smoke (Goniewicz et al., 2014), thus switching completely from smoking to vaping should reduce their harm (Shahab, Goniewicz, et al., 2017).

The data collected so far in the experimental study described in Chapter 6 are in line with the hypothesis that nicotine decreases food consumption rate but are not in line with the hypothesis that nicotine decreases food consumption or self-selected portion size. I cannot yet derive any clear conclusions from this data, as the sample size is smaller than calculated to achieve sufficient statistical power, but the methods appear to be robust enough to draw conclusions when the pre-determined sample size is achieved. If appetite is affected by nicotine exposure as hypothesised, nicotine containing e-cigarette use could be encouraged among smokers who have relapsed due to post-cessation weight gain (Mizes et al., 1998) to potentially increase the likelihood of a successful quit attempt. If appetite is not affected, further investigation will be needed to identify potential mechanisms through which nicotine exposure and BMI are related.

7.3 Thesis Strengths

7.3.1 Novel and Improved Methods

A strength of this thesis is that I have extended previously used methods and used novel methods to address important questions regarding the factors influencing e-cigarette use and the possible consequences of their use.

In Chapter 2, I updated a previous meta-analysis exploring whether e-cigarettes act as a gateway to smoking (Soneji, Barrington-Trimis, Wills, Leventhal, et al., 2017) by adding relevant search terms to the search criteria, providing further details of the included studies, and stratifying the data to explore sources of heterogeneity. The previous meta-analysis also appears to include an error; the authors stated that they included only studies reporting odds ratios, yet Miech and colleagues (2017) reported relative risk. When contacted, the corresponding author stated that the relative risk had been converted, yet the converted estimate was identical to the original estimate of relative

risk. The author did not respond to a request to share their equation for conversion. It is possible that an odds ratio and relative risk would be similar if the outcome is rare (Davies, Crombie, & Tavakoli, 1998), but in my conversion the estimate was not equivalent to the relative risk reported. Including a relative risk in a meta-analysis with odds ratios would be an issue for the interpretation of the combined estimate as interpreting relative risks as odds ratios will understate the effect (Davies et al., 1998). Therefore, I have provided an update to the previous meta-analysis with which the combined estimate can be more easily interpreted.

In Chapter 3, I explored why young adults in the UK use e-cigarettes. Young adults are an understudied population – particularly in the UK – because they are often included in studies with older adults, assuming that the young adults and older adults share the same reasons for vaping. I have shown that this is not the case; the most common reason for vaping among young adults (23 years old) in the UK was out of curiosity, whereas among adults in Great Britain (18+ years old) the most common reason is to quit smoking (Action on Smoking and Health, 2019a).

In Chapter 4, I extended the methods of Allegrini and colleagues (2019) by using updated smoking initiation polygenic risk scores. The updated scores were created using the summary results of a larger genome-wide association study with greater statistical power to detect significant genetic variants (Liu et al., 2019). As a result, the scores I used should explain more variance in smoking initiation. My method was also novel as I included negative control outcomes to explore potential horizontal pleiotropy and shared liability.

In Chapter 5, I employed a novel method to explore the potential health effects of nicotine use without exposure to cigarette smoke. To my knowledge, this is the first study to have used multivariable Mendelian randomisation to investigate and separate these effects. Without novel methods such as multivariable Mendelian randomisation, it would be decades before these effects would be observed, and the observed measures would likely be biased by confounding factors. Applying this novel multivariable Mendelian randomisation approach to other complex behaviours could be useful to separate the effects of specific components of the behaviour.

In Chapter 6, I used novel methods to explore the effects of nicotine on appetite by recruiting non-dependent smokers. Previous studies (Perkins et al., 1992; Perkins et al., 1991) have recruited non-smokers (who may experience adverse effects due to being naïve to nicotine) and daily smokers (who may experience withdrawal symptoms which could affect their appetite). Non-dependent smokers have been excluded from previous studies but exploring the effects among individuals who are unlikely to experience nicotine withdrawal or adverse effects due to naivety to nicotine could help to explain the role of nicotine in lowering BMI among smokers.

7.3.2 Sample Size and Statistical Power

A second strength of this thesis is that I have used relatively large sample sizes, which increases statistical power to detect true effects (Button et al., 2013). I have described the largest meta-analysis to date which explores whether e-cigarette use acts as a gateway to smoking, including 17 studies with sample sizes ranging from 347 to 39,718 (Chapter 2). I have also utilised large, prospective cohorts with rich data (the Avon Longitudinal Study of Parents and Children [N = 412 to 6,702] and UK Biobank [N = 337,010]) and used summary genetic data from GSCAN (Liu et al., 2019), a large-scale genome-wide association study with a sample size of 337,334 (Chapters 3, 4, and 5). Furthermore, using the Avon Longitudinal Study of Parents and Children data set, I was able to access a large sample of young adults which allowed me to investigate my research questions among an understudied subgroup of individuals. However, large sample sizes are required for genetic studies and so I may still have been underpowered to detect effects in Chapters 4 and 5.

I also calculated the appropriate sample size to achieve 90% power to detect a substantial effect in the experiment described in Chapter 6. However, as this sample size has not yet been achieved, I was not able to report any inferential statistics and so I cannot be confident in the magnitude or direction of the effects found thus far.

7.3.3 Causal Inference and Strength of Evidence

A third strength of this thesis is that I have used a range of methods which are considered to contribute strong evidence of causation and aid consideration of possible causal pathways. For example, in Chapter 4 the use of negative controls aided my

consideration of whether the relationship between smoking initiation polygenic risk scores and e-cigarette use is causal (i.e., the genetic variants influence e-cigarette use only through smoking). By including negative controls relating to risk-taking in childhood (before the vast majority of participants engaged in smoking), I was able to show that there could be an underlying risk-taking phenotype which cannot be a result of smoking (as smoking had not occurred at this age) and which could also influence e-cigarette use.

Traditionally, randomised controlled trials and meta-analyses are stated to be the strongest forms of evidence (Haapasalo, 2016). Mendelian randomisation is akin to a randomised controlled trial due to the random assortment of genetic variants at conception (Nitsch et al., 2006; Swanson et al., 2017). Although I have included strong methods in this thesis (i.e., meta-analyses and Mendelian randomisation), such methods can still be limited by bias and error. By triangulating these methods and drawing similar conclusions regarding the aims of my thesis, I provide stronger support for my conclusions than if I had presented the evidence from any of my individual studies alone (Munafo & Davey Smith, 2018).

7.4 Thesis Limitations

7.4.1 Risk of Bias, Measurement Error and Reverse Causality

The first limitation of this thesis is that risk of bias, measurement error and potential reverse causality may have influenced some of my findings. Including studies of low quality (i.e., studies with greater potential for bias) in meta-analyses can lead to an issue referred to as 'garbage in, garbage out' whereby the results of the meta-analysis are also biased (Sharpe, 1997). Although I assessed and stratified by risk of bias in my meta-analysis in Chapter 2, the criteria which I used to determine the quality of the studies was relatively liberal as there were so few studies which met stricter criteria for high quality, as highlighted in a recent review focussing on adolescents (Chan et al., 2020). For example, using a star-based rating system, I awarded a star to studies which had used self-reports to demonstrate that the outcome of interest preceded the exposure, but smoking status is sometimes misreported by young people (Khouja, Munafò, et al., 2020), so the outcome of interest may not have preceded the exposure in all cases. If I had only awarded stars to studies using biochemical verification to

determine the order of exposure and outcome (thereby removing some measurement error and limiting potential reverse causality), no stars would have been awarded for this criterion. Consequently, all of the studies would have been rated poor/fair quality and no studies would have been rated as good quality. Although it is possible to biochemically verify whether non-current smokers are currently vaping (by verifying nicotine exposure without carbon monoxide exposure), it is not currently possible to biochemically verify that a person has ever vaped but never smoked. However, this may be possible in the near future with advancing methods such as DNA methylation (Richmond et al., 2018), therefore more research using stronger emerging methods is needed.

Measurement error and reverse causality may also have affected the results of Chapter 3, as the young adults retrospectively-recalled their reasons for using e-cigarettes. This may lead to recall bias; for example, if an individual used an e-cigarette to quit smoking but failed to quit, they may be less likely to recall that they used an e-cigarette to quit. Ideally, I would have measured reasons for use just prior to vaping for the first time but this may also influence my results. For example, I could recruit individuals who intend to vape but have not yet tried vaping, but this would exclude individuals who try e-cigarettes on an impulse which could substantially reduce the proportion of vapers who vaped out of curiosity included in the study.

7.4.2 Data Availability and Weak-Instrument Bias

The second limitation of this thesis is that my results may have been restricted by the data which was available to me. For example, ALSPAC is a rich source of data, but may not be appropriate for studying the gateway effect given the context of the cohort (i.e., the young adults were exposed to cigarettes but not e-cigarettes in their youth).

Due to only one small GWAS of cotinine being available (and no GWAS of e-cigarette use) my results may also have been influenced by weak-instrument bias. In Chapter 5, the conditional F-statistic for multivariable Mendelian randomisation indicated that there may be weak-instrument bias in the multivariable Mendelian randomisation analysis of the direct effect of nicotine. However, I was able to investigate the effect of nicotine by observing differences (or lack thereof) between univariable and multivariable Mendelian randomisation of smoking heaviness. As there was no clear

evidence of weak-instrument bias in the analysis of the direct effect of smoking heaviness, a difference between the direct and total effects provides some evidence of an indirect effect via the other exposure(s) accounted for in the multivariable Mendelian randomisation analyses (i.e., nicotine). Therefore, comparing the univariable and multivariable Mendelian randomisation analyses of smoking heaviness and smoking-related health outcomes allowed me to make cautious inferences about the potential direct effects of nicotine on smoking-related health outcomes.

7.4.3 Horizontal Pleiotropy and Collider Bias

Aside from weak instrument bias, Mendelian randomisation studies can also be biased when the assumptions of Mendelian randomisation are violated. For example, results will be biased when the genetic variants included in a model are not valid because they exert horizontally pleiotropic effects (i.e., when one genetic variant affects more than one different phenotype) (Davies et al., 2018). Conditioning or stratifying on a collider can also bias the results of Mendelian randomisation (Davies et al., 2018) and observational studies (Munafò et al., 2018). The third limitation of this thesis is that horizontal pleiotropy and collider bias may have occurred.

In Chapter 4, I found evidence to suggest that the genetic variants used to create the smoking initiation polygenic risk scores could have been horizontally pleiotropic as they were not only associated with smoking and e-cigarette use, but also with risk-taking, impulsivity and socioeconomic position at birth. Pleiotropy robust methods such as MR-Egger have been developed to address issues of pleiotropy, but MR-Egger allows for pleiotropy only if the InSIDE (Instrument Strength Independent of Direct Effect) assumption is held. These associations demonstrate a violation of the InSIDE assumption, so even pleiotropy robust methods, such as MR-Egger, may not be appropriate when conducting Mendelian randomisation studies of smoking initiation. In Chapter 5, I instead conducted univariable and multivariable Mendelian randomisation analyses of smoking heaviness but there was also evidence of horizontal pleiotropy in these analyses (as indicated by the MR-Egger intercept, Cochran's Q statistic, and effects observed among never smokers).

The effects seen among never smokers may alternatively be due to stratifying on the exposure, which may introduce collider bias. I stratified the analyses by smoking status

(i.e., never, ever, former and current smoking), which is likely to be impacted by the number of cigarettes an individual is predisposed to smoke per day, and health status. The smoking heaviness genetic variant rs16969968-rs1051730 has previously been suggested to be positively associated with smoking status (i.e., ever versus never smoking) in younger age groups but negatively associated with smoking status in older age groups – smokers with the heavier smoking genetic predisposition have a higher mortality rate and are therefore less likely to survive into older adulthood (Taylor & Munafò, 2014). Diagnosis of a smoking-related disease could also lead smokers to quit smoking; health concern is a particularly strong motivation to quit smoking among those with a diagnosed disease (Buczkowski, Marcinowicz, Czachowski, & Piszczek, 2014). As both the exposure and outcome could influence smoking status, smoking status is a collider and thus stratifying by it could induce collider bias. Similarly, selection into UK Biobank may have introduced collider bias as participants are less likely to smoke than the general population and the mortality rate is also lower (Munafò et al., 2018). Consequently, the Mendelian randomisation results reported in Chapter 5 may be subject to bias whereby associations are negatively biased (i.e., results may indicate stronger evidence that increased smoking heaviness causes poor health than the true effect).

The analysis in Chapter 4 may also have been subject to collider bias. If both smoking initiation polygenic risk scores and unmeasured confounders influence smoking initiation (i.e., ever versus never smoking), then smoking initiation is a collider in this analysis. Thus conditioning on smoking initiation (i.e., restricting the analysis to never smokers) as I have in this analysis could induce bias (Paternoster et al., 2017) which could lead to biased associations in either direction depending on the direction of the effect of the unmeasured confounders.

7.5 Future Directions

Given the limitations I have outlined, further research is needed to develop our understanding of the factors influencing youth and young adult e-cigarette use and the potential consequences of use.

First, the evidence base exploring whether e-cigarette use acts as a gateway to smoking is rapidly expanding meaning that my systematic review and meta-analysis could soon be out of date. Consequently, frequent updates of this review are necessary and should take into account advances and developments in the field. For instance, if more studies account for the nicotine content and frequency of exposure, include biochemical verification and adjust more appropriately for confounders, the results of the meta-analysis may differ from the findings presented in Chapter 2. Stricter criteria could also be used to determine the quality of studies, in which case subsequent stratification may result in different findings. However, if future studies also fail to take these factors into account, the results of future meta-analyses are likely to result in similar conclusions as the evidence of an association between e-cigarette use and later smoking appears to be robust.

Second, with new and emerging data, the procedures described in this thesis could be replicated in other relevant populations or adapted to more thoroughly explore the research questions. For example, the Millennium cohort is a prospective birth cohort study which released genetic data in Autumn 2020 – the young people in this study were exposed equally to e-cigarettes and cigarettes in youth and thus would be an appropriate population to explore whether smoking initiation polygenic risk scores predict e-cigarette use among never-smoking youth. Replicating the findings of Chapter 4 in this cohort (or other similar cohorts) could provide further support to my conclusion that the presumed gateway effect is better explained as a shared liability. Additionally, with the greater availability of cohorts with genetic and phenotypic data, a genome-wide association study of e-cigarette use may become available which could be used to directly compare the genetic influences on vaping with the genetic influences on smoking (furthering the findings of Chapter 4) and explore causal effects of e-cigarette use (furthering the findings of Chapters 2 and 5). Following on from my PhD, I plan to form a consortium aimed at collating genetic and phenotypic data relating to e-cigarette use from a range of cohorts to facilitate a genome-wide association study of e-cigarette use.

Third, I plan to resume recruitment for the study exploring the effect of nicotine on eating behaviour (described in Chapter 6) and will continue collecting data until the pre-planned sample size is achieved. However, I may need to adapt some of the

procedure in order to meet new government guidelines for safety during (or after) the COVID-19 pandemic. During the initial recruitment phase (pre-pandemic), the methods were refined in order to maximise recruitment, drop-out rates were moderate, and the data (bar data measured using the Mandometer) were of good quality, thus I am confident that with a sufficient sample size the data will be adequate to proceed with my planned analyses. If the results support my hypotheses (i.e., nicotine administration decreases food consumption, self-selected portion size and consumption rate compared with placebo) I will conclude that nicotine appears to reduce appetite among non-dependent smokers and this could explain the observed relationship between smoking and lower BMI. If the data do not support my hypotheses, I will conclude that this relationship may not be due to a nicotine-induced reduction in appetite and food consumption.

7.6 Public Health and Policy Implications

Despite the limitations of this thesis, my findings have substantial implications for policy and public health strategies relating to e-cigarettes, particularly for strategies aimed at reducing harm to youth and young adults. The findings in Chapter 2 do not provide strong support for the gateway hypothesis yet the combined findings of Chapters 2, 3 and 4 provide strong support for a shared liability between vaping and smoking, therefore, strict policies (e.g., bans) which prevent e-cigarette use in order to reduce the risk of smoking initiation among youth and young adults are unlikely to be effective. In fact, they may have the opposite effect; if young people are predisposed to both vaping and smoking but only cigarettes are available, this could increase their likelihood of smoking because it is the only option available to them.

Given the strong positive association that I reported between e-cigarette use among never-smokers and later smoking in Chapter 2, researchers and policymakers alike are interested in what may attract non-smokers to e-cigarettes in order to make them unattractive. Frequently, attractive flavours and packaging have been identified as potential factors which could be attracting young non-smokers, and as a result, restrictions have been placed on flavour availability and packaging in some countries (Klein, Chaiton, Kundu, & Schwartz, 2020). However, there is limited evidence of the impact that flavours have on e-cigarette use among non-smokers, and there is very

limited evidence of the impact that flavour restrictions would have on current smokers and vapers. Some evidence suggests that some vapers would relapse to smoking and some would create their own flavours or source their e-liquids from unregulated/illegal sources, which can have lethal consequences such as E-cigarette and Vaping Associated Lung Injury (Action on Smoking and Health, 2020; Heinzerling et al., 2020). Thus, restrictions on flavours may actually put current vapers at risk and given that the association found between vaping and later smoking in Chapter 2 may not be causal, such restrictions may not protect non-smokers or vapers.

Furthermore, restrictive policies may prevent and discourage adult smokers from accessing an effective smoking cessation tool and hamper smoking cessation attempts. The evidence I presented in Chapter 5 suggests that consuming nicotine without tobacco smoke (e.g., via e-cigarettes rather than cigarettes) is likely to reduce the risk of developing smoking-related diseases. Therefore, policies should encourage smokers to switch from tobacco cigarettes to e-cigarettes. Additionally, restrictions placed on nicotine contents (such as the 20 mg/ml restriction in the UK) should be carefully considered to ensure that heavy smokers have access to a less harmful nicotine replacement which is sufficient for their needs.

The results I have presented suggest that policies which encourage smokers to switch to e-cigarettes should be adopted over those that discourage use. Ideally, policy should encourage smokers to vape while simultaneously restricting non-smokers' access to both cigarettes and e-cigarettes.

7.7 Thesis Conclusion

In this thesis I aimed to explore who uses e-cigarettes, why, and what the potential consequences of e-cigarette use could be, particularly among young adults. I highlighted the need to address these aims, showing that young people who have never smoked before are being exposed to vaping, which *may* act as a gateway to smoking. I found substantial differences in demographic, behavioural and lifestyle factors between young adults who had ever and never vaped, as well as differences between those who were former and current vapers. Specifically, vapers were more likely to be male, of lower socioeconomic position at birth, have poorer mental health, engage in other risky

behaviours, and smoke. I also identified curiosity as the most popular reason for using e-cigarettes (followed by use to quit smoking) among a sample of UK young adults who had both ever smoked and ever vaped. Vaping out of curiosity did not appear to change an individual's smoking or vaping status (i.e., individuals were unlikely to become current vapers and smokers continued to smoke) whereas vaping to quit smoking was associated with continued vaping and discontinued smoking. I also showed that e-cigarette use may be influenced by a genetic predisposition to risk-taking. Finally, I showed that nicotine exposure (without exposure to cigarette smoke) does not appear to cause smoking-related ill health.

Using a variety of methods, I have shown that the shared liability hypothesis may explain the relationship between e-cigarette use among non-smokers and later smoking, and that nicotine use without exposure to tobacco smoke (e.g., via e-cigarettes) is unlikely to cause smoking-related ill health. However, further research is required to support these findings using emerging datasets which will allow for greater exploration of the gateway and shared liability hypotheses as well as potential consequences of vaping.

References

Action on Smoking and Health. (2017). *Use of e-cigarettes (vapourisers) among adults in Great Britain*. Retrieved from file:///C:/Users/jk14427/Chrome%20Local%20Downloads/Use-of-e-cigarettes-vapourisers-among-adults-in-Great-Britain-May-2017-2%20(3).pdf

Action on Smoking and Health. (2019a). *Use of e-cigarettes (vapourisers) among adults in Great Britain*. Retrieved from https://ash.org.uk/wp-content/uploads/2019/09/Use-of-e-cigarettes-among-adults-2019.pdf

Action on Smoking and Health. (2019b). *Use of e-cigarettes among young people in Great Britain*. Retrieved from https://ash.org.uk/wp-content/uploads/2019/06/ASH-Factsheet-Youth-E-cigarette-Use-2019.pdf

Action on Smoking and Health. (2020). *Use of e-cigarettes (vapes) among adults in Great Britain*. Retrieved from https://ash.org.uk/wp-content/uploads/2020/10/Use-of-e-cigarettes-vapes-among-adults-in-Great-Britain-2020.pdf

Adkison, S. E., O'Connor, R. J., Bansal-Travers, M., Hyland, A., Borland, R., Yong, H. H., . . . Fong, G. T. (2013). Electronic Nicotine Delivery Systems: International Tobacco Control Four-Country Survey. *Am J Prev Med, 44*(3), 207-215. doi:10.1016/j.amepre.2012.10.018

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). *The Structure and Function of DNA*: Garland Science.

Aleyan, S., Hitchman, S. C., Ferro, M. A., & Leatherdale, S. T. (2020). Trends and predictors of exclusive e-cigarette use, exclusive smoking and dual use among youth in Canada. *Addictive Behaviors, 109*, 106481. doi:https://doi.org/10.1016/j.addbeh.2020.106481

Allegrini, A. G., Verweij, K. J. H., Abdellaoui, A., Treur, J. L., Hottenga, J. J., Willemsen, G., . . . Vink, J. M. (2019). Genetic Vulnerability for Smoking and Cannabis Use: Associations With E-Cigarette and Water Pipe Use. *Nicotine Tob Res, 21*(6), 723-730. doi:10.1093/ntr/nty150

Allen, J. G., Flanigan, S. S., LeBlanc, M., Vallarino, J., MacNaughton, P., Stewart, J. H., & Christiani, D. C. (2016). Flavoring Chemicals in E-Cigarettes: Diacetyl, 2,3-Pentanedione, and Acetoin in a Sample of 51 Products, Including Fruit-, Candy-, and Cocktail-Flavored E-Cigarettes. *Environ Health Perspect, 124*(6), 733-739. doi:10.1289/ehp.1510185

Allen, N. E., Sudlow, C., Peakman, T., Collins, R., & Biobank, U. K. (2014). UK biobank data: come and get it. *Sci Transl Med, 6*(224), 224ed224. doi:10.1126/scitranslmed.3008601

Åsvold, B. O., Bjørngaard, J. H., Carslake, D., Gabrielsen, M. E., Skorpen, F., Davey Smith, G., & Romundstad, P. R. (2014). Causal associations of tobacco smoking with

cardiovascular risk factors: a Mendelian randomization analysis of the HUNT Study in Norway. *International Journal of Epidemiology, 43*(5), 1458-1470. doi:10.1093/ije/dyu113

Audrain-McGovern, J., & Benowitz, N. L. (2011). Cigarette smoking, nicotine, and body weight. *Clin Pharmacol Ther, 90*(1), 164-168. doi:10.1038/clpt.2011.105

Auf, R., Trepka, M. J., Selim, M., Ben Taleb, Z., De La Rosa, M., Bastida, E., & Cano, M. A. (2018). E-cigarette use is associated with other tobacco use among US adolescents. *Int J Public Health*. doi:10.1007/s00038-018-1166-7

Barrington-Trimis, J. L., Kong, G., Leventhal, A. M., Liu, F., Mayer, M., Cruz, T. B., . . . McConnell, R. (2018). E-cigarette Use and Subsequent Smoking Frequency Among Adolescents. *Pediatrics*. doi:10.1542/peds.2018-0486

Barrington-Trimis, J. L., Urman, R., Berhane, K., Unger, J. B., Cruz, T. B., Pentz, M. A., . . . McConnell, R. (2016). E-Cigarettes and Future Cigarette Use. *Pediatrics, 138*(1). doi:10.1542/peds.2016-0379

Bauld, L., MacKintosh, A. M., Eastwood, B., Ford, A., Moore, G., Dockrell, M., . . . McNeill, A. (2017). Young People's Use of E-Cigarettes across the United Kingdom: Findings from Five Surveys 2015-2017. *Int J Environ Res Public Health, 14*(9). doi:10.3390/ijerph14090973

Behar, R. Z., Luo, W., Lin, S. C., Wang, Y., Valle, J., Pankow, J. F., & Talbot, P. (2016). Distribution, quantification and toxicity of cinnamaldehyde in electronic cigarette refill fluids and aerosols. *Tob Control, 25*(Suppl 2), ii94-ii102. doi:10.1136/tobaccocontrol-2016-053224

Bell, K., & Keane, H. (2014). All gates lead to smoking: The 'gateway theory', e-cigarettes and the remaking of nicotine. *Social Science & Medicine, 119*, 45-52. doi:10.1016/j.socscimed.2014.08.016

Benowitz, N. L. (2010). Nicotine Addiction. *New England Journal of Medicine, 362*(24), 2295-2303. doi:10.1056/NEJMra0809890

Benowitz, N. L., & Fraiman, J. B. (2017). Cardiovascular effects of electronic cigarettes. *Nature Reviews Cardiology, 14*(8), 447-456. doi:10.1038/nrcardio.2017.36

Benowitz, N. L., & Gourlay, S. G. (1997). Cardiovascular Toxicity of Nicotine: Implications for Nicotine Replacement Therapy 11All editorial decisions for this article, including selection of referees, were made by a Guest Editor. This policy applies to all articles with authors from the University of California San Francisco. *Journal of the American College of Cardiology, 29*(7), 1422-1431. doi:https://doi.org/10.1016/S0735-1097(97)00079-X

Benowitz, N. L., Hukkanen, J., & Jacob, P., 3rd. (2009). Nicotine chemistry, metabolism, kinetics and biomarkers. *Handb Exp Pharmacol*(192), 29-60. doi:10.1007/978-3-540-69248-5_2

References

Berlin, I., Dautzenberg, B., Lehmann, B., Palmyre, J., Liégey, E., De Rycke, Y., & Tubach, F. (2019). Randomised, placebo-controlled, double-blind, double-dummy, multicentre trial comparing electronic cigarettes with nicotine to varenicline and to electronic cigarettes without nicotine: the ECSMOKE trial protocol. *BMJ Open, 9*(5), e028832. doi:10.1136/bmjopen-2018-028832

Besaratinia, A., & Tommasi, S. (2019). Vaping: A growing global health concern. *EClinicalMedicine, 17*, 100208. doi:10.1016/j.eclinm.2019.10.019

Best, C., Haseen, F., Currie, D., Ozakinci, G., MacKintosh, A. M., Stead, M., . . . Haw, S. (2018). Relationship between trying an electronic cigarette and subsequent cigarette experimentation in Scottish adolescents: a cohort study. *Tobacco Control, 27*(4), 373-378. doi:10.1136/tobaccocontrol-2017-053691

Bhatta, D. N., & Glantz, S. A. (2019). Electronic Cigarette Use and Myocardial Infarction Among Adults in the US Population Assessment of Tobacco and Health. *J Am Heart Assoc, 8*(12), e012317. doi:10.1161/jaha.119.012317

Biener, L., Song, E., Sutfin, E. L., Spangler, J., & Wolfson, M. (2015). Electronic Cigarette Trial and Use among Young Adults: Reasons for Trial and Cessation of Vaping. *International Journal of Environmental Research and Public Health, 12*(12), 16019-16026. doi:10.3390/ijerph121215039

Blount, B. C., Karwowski, M. P., Shields, P. G., Morel-Espinosa, M., Valentin-Blasini, L., Gardner, M., . . . Pirkle, J. L. (2019). Vitamin E Acetate in Bronchoalveolar-Lavage Fluid Associated with EVALI. *New England Journal of Medicine, 382*(8), 697-705. doi:10.1056/NEJMoa1916433

Bold, K. W., Kong, G., Camenga, D. R., Simon, P., Cavallo, D. A., Morean, M. E., & Krishnan-Sarin, S. (2018). Trajectories of E-Cigarette and Conventional Cigarette Use Among Youth. *Pediatrics, 141*(1), e20171832. doi:10.1542/peds.2017-1832

Bold, K. W., Kong, G., Cavallo, D. A., Camenga, D. R., & Krishnan-Sarin, S. (2016). Reasons for Trying E-cigarettes and Risk of Continued Use. *Pediatrics, 138*(3). doi:10.1542/peds.2016-0895

Bowden, J., Davey Smith, G., & Burgess, S. (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol, 44*(2), 512-525. doi:10.1093/ije/dyv080

Bowden, J., Davey Smith, G., Haycock, P. C., & Burgess, S. (2016). Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol, 40*(4), 304-314. doi:10.1002/gepi.21965

Bowden, J., Del Greco, M. F., Minelli, C., Davey Smith, G., Sheehan, N. A., & Thompson, J. R. (2016). Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I2 statistic. *Int J Epidemiol, 45*(6), 1961-1974. doi:10.1093/ije/dyw220

Bowden, J., & Holmes, M. V. (2019). Meta-analysis and Mendelian randomization: A review. *Res Synth Methods, 10*(4), 486-496. doi:10.1002/jrsm.1346

References

Bowen, A., & Xing, C. (2015) Nicotine salt formulations for aerosol devices and methods thereof. U.S. Patent No. 9,215,895. U. S. P. a. T. Office.

Boyd, A., Golding, J., Macleod, J., Lawlor, D. A., Fraser, A., Henderson, J., . . . Smith, G. D. (2013). Cohort Profile: The 'Children of the 90s'-the index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology, 42*(1), 111-127. doi:10.1093/ije/dys064

Brown, J., West, R., Beard, E., Michie, S., Shahab, L., & McNeill, A. (2014). Prevalence and characteristics of e-cigarette users in Great Britain: Findings from a general population survey of smokers. *Addict Behav, 39*(6), 1120-1125. doi:10.1016/j.addbeh.2014.03.009

Brown, R., Bauld, L., de Lacy, E., Hallingberg, B., Maynard, O., McKell, J., . . . Moore, G. (2020). A qualitative study of e-cigarette emergence and the potential for renormalisation of smoking in UK youth. *International Journal of Drug Policy, 75*, 102598. doi:https://doi.org/10.1016/j.drugpo.2019.11.006

Buczkowski, K., Marcinowicz, L., Czachowski, S., & Piszczek, E. (2014). Motivations toward smoking cessation, reasons for relapse, and modes of quitting: results from a qualitative study among former and current smokers. *Patient Prefer Adherence, 8*, 1353-1363. doi:10.2147/ppa.s67767

Burgess, S., Butterworth, A., & Thompson, S. G. (2013). Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol, 37*(7), 658-665. doi:10.1002/gepi.21758

Bush, W. S., & Moore, J. H. (2012). Chapter 11: Genome-Wide Association Studies. *PLOS Computational Biology, 8*(12), e1002822. doi:10.1371/journal.pcbi.1002822

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafo, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci, 14*(5), 365-376. doi:10.1038/nrn3475

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., . . . Marchini, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature, 562*(7726), 203-209. doi:10.1038/s41586-018-0579-z

CDC. (2020). Outbreak of Lung Injury Associated with the Use of E-Cigarette, or Vaping, Products. Retrieved from https://www.cdc.gov/tobacco/basic_information/e-cigarettes/severe-lung-disease.html#key-facts-vit-e

Chakma, J. K., Kumar, H., Bhargava, S., & Khanna, T. (2020). The e-cigarettes ban in India: an important public health decision. *The Lancet Public Health, 5*(8), e426. doi:10.1016/S2468-2667(20)30063-3

Chan, G. C. K., Stjepanović, D., Lim, C., Sun, T., Shanmuga Anandan, A., Connor, J. P., . . . Leung, J. (2020). Gateway or common liability? A systematic review and meta-analysis of studies of adolescent e-cigarette use and future smoking initiation. *Addiction, n/a*(n/a). doi:10.1111/add.15246

References

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience, 4*, 7. doi:10.1186/s13742-015-0047-8

Chapman, S., Bareham, D., & Maziak, W. (2019). The Gateway Effect of E-cigarettes: Reflections on Main Criticisms. *Nicotine Tob Res, 21*(5), 695-698. doi:10.1093/ntr/nty067

Choi, S. W., & O'Reilly, P. F. (2019). PRSice-2: Polygenic Risk Score software for biobank-scale data. *Gigascience, 8*(7). doi:10.1093/gigascience/giz082

Cole, S. R., Platt, R. W., Schisterman, E. F., Chu, H. T., Westreich, D., Richardson, D., & Poole, C. (2010). Illustrating bias due to conditioning on a collider. *International Journal of Epidemiology, 39*(2), 417-420. doi:10.1093/ije/dyp334

Conner, M., Grogan, S., Simms-Ellis, R., Flett, K., Sykes-Muskett, B., Cowap, L., . . . Siddiqi, K. (2018). Do electronic cigarettes increase cigarette smoking in UK adolescents? Evidence from a 12-month prospective study. *Tobacco Control, 27*(4), 365-372. doi:10.1136/tobaccocontrol-2016-053539

Das, S. K. (2003). Harmful health effects of cigarette smoking. *Mol Cell Biochem, 253*(1-2), 159-165. doi:10.1023/a:1026024829294

Davey Smith, G., & Ebrahim, S. (2003). 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?*. *International Journal of Epidemiology, 32*(1), 1-22. doi:10.1093/ije/dyg070

Davey Smith, G., & Hemani, G. (2014). Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet, 23*(R1), R89-98. doi:10.1093/hmg/ddu328

Davey Smith, G., Phillips, A. N., & Neaton, J. D. (1992). Smoking as Independent Risk Factor for Suicide - Illustration of an Artifact from Observational Epidemiology. *Lancet, 340*(8821), 709-712. Retrieved from <Go to ISI>://WOS:A1992JN78000014

Davies, H. T., Crombie, I. K., & Tavakoli, M. (1998). When can odds ratios mislead? *BMJ (Clinical research ed.), 316*(7136), 989-991. doi:10.1136/bmj.316.7136.989

Davies, N. M., Holmes, M. V., & Davey Smith, G. (2018). Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ, 362*, k601. doi:10.1136/bmj.k601

DeVito, E. E., & Krishnan-Sarin, S. (2018). E-cigarettes: Impact of E-Liquid Components and Device Characteristics on Nicotine Exposure. *Current neuropharmacology, 16*(4), 438-459. doi:10.2174/1570159X15666171016164430

Domoff, S. E. (2015). Dutch Eating Behavior Questionnaire (DEBQ). In T. Wade (Ed.), *Encyclopedia of Feeding and Eating Disorders* (pp. 1-2). Singapore: Springer Singapore.

References

East, K., Hitchman, S. C., Bakolis, I., Williams, S., Cheeseman, H., Arnott, D., & McNeill, A. (2018). The Association Between Smoking and Electronic Cigarette Use in a Cohort of Young People. *Journal of Adolescent Health, 62*(5), 539-547. doi:10.1016/j.jadohealth.2017.11.301

Eastwood, B., Dockrell, M. J., Arnott, D., Britton, J., Cheeseman, H., Jarvis, M. J., & McNeill, A. (2015). Electronic cigarette use in young people in Great Britain 2013-2014. *Public Health, 129*(9), 1150-1156. doi:10.1016/j.puhe.2015.07.009

Etter, J. F. (2018). Gateway effects and electronic cigarettes. *Addiction, 113*(10), 1776-1783. doi:10.1111/add.13924

Euesden, J., Lewis, C. M., & O'Reilly, P. F. (2015). PRSice: Polygenic Risk Score software. *Bioinformatics, 31*(9), 1466-1468. doi:10.1093/bioinformatics/btu848

Evans, D. L. (1985). Cardiovascular adaptations to exercise and training. *Vet Clin North Am Equine Pract, 1*(3), 513-531. doi:10.1016/s0749-0739(17)30748-4

Fairchild, A. L., Bayer, R., & Colgrove, J. (2013). The Renormalization of Smoking? E-Cigarettes and the Tobacco "Endgame". *New England Journal of Medicine, 370*(4), 293-295. doi:10.1056/NEJMp1313940

Farsalinos, K. E., Spyrou, A., Tsimopoulou, K., Stefopoulos, C., Romagna, G., & Voudris, V. (2014). Nicotine absorption from electronic cigarette use: comparison between first and new-generation devices. *Scientific Reports, 4*. doi:10.1038/srep04133

Farsalinos, K. E., Voudris, V., & Poulas, K. (2015). E-cigarettes generate high levels of aldehydes only in 'dry puff' conditions. *Addiction, 110*(8), 1352-1356. doi:10.1111/add.12942

Field, A. (2013). *Discovering Statistics using IBM SPSS Statistics*: Sage Publications Ltd.

Filippidis, F. T., Laverty, A. A., Gerovasili, V., & Vardavas, C. I. (2017). Two-year trends and predictors of e-cigarette use in 27 European Union member states. *Tobacco Control, 26*(1), 98-104. doi:10.1136/tobaccocontrol-2015-052771

Ford, A. L., Bergh, C., Sodersten, P., Sabin, M. A., Hollinghurst, S., Hunt, L. P., & Shield, J. P. (2009). Treatment of childhood obesity by retraining eating behaviour: randomised controlled trial. *BMJ, 340*, b5388. doi:10.1136/bmj.b5388

Fraser, A., Macdonald-Wallis, C., Tilling, K., Boyd, A., Golding, J., Davey Smith, G., . . . Lawlor, D. A. (2013). Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol, 42*(1), 97-110. doi:10.1093/ije/dys066

Fraser, A., Macdonald-Wallis, C., Tilling, K., Boyd, A., Golding, J., Smith, G. D., . . . Lawlor, D. A. (2013). Cohort Profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *International Journal of Epidemiology, 42*(1), 97-110. doi:10.1093/ije/dys066

References

Gage, S. H., Davey Smith, G., Ware, J. J., Flint, J., & Munafo, M. R. (2016). G = E: What GWAS Can Tell Us about the Environment. *PLoS Genet, 12*(2), e1005765. doi:10.1371/journal.pgen.1005765

Galobardes, B., Shaw, M., Lawlor, D. A., & Lynch, J. W. (2006a). Indicators of socioeconomic position (part 1). In *J Epidemiol Community Health* (Vol. 60, pp. 7-12).

Galobardes, B., Shaw, M., Lawlor, D. A., & Lynch, J. W. (2006b). Indicators of socioeconomic position (part 2). In *J Epidemiol Community Health* (Vol. 60, pp. 95-101).

Gartner, C., & Bromberg, M. (2019). The Regulation of E-cigarettes. In *One does not simply sell e-cigarettes in Australia: an overview of Australian e-cigarette regulations*: Edward Elgar Publishing.

Gholap, V. V., Heyder, R. S., Kosmider, L., & Halquist, M. S. (2020). An Analytical Perspective on Determination of Free Base Nicotine in E-Liquids. *J Anal Methods Chem, 2020*, 6178570. doi:10.1155/2020/6178570

Gillman, I. G., Kistler, K. A., Stewart, E. W., & Paolantonio, A. R. (2016). Effect of variable power levels on the yield of total aerosol mass and formation of aldehydes in e-cigarette aerosols. *Regulatory Toxicology and Pharmacology, 75*, 58-65. doi:https://doi.org/10.1016/j.yrtph.2015.12.019

Glantz, S. A., & Bareham, D. W. (2018). E-Cigarettes: Use, Effects on Smoking, Risks, and Policy Implications. *Annu Rev Public Health, 39*, 215-235. doi:10.1146/annurev-publhealth-040617-013757

Glasser, A., Abudayyeh, H., Cantrell, J., & Niaura, R. (2018). Patterns of E-Cigarette Use Among Youth and Young Adults: Review of the Impact of E-Cigarettes on Cigarette Smoking. *Nicotine Tob Res*. doi:10.1093/ntr/nty103

Godish, T. (1989). Formaldehyde exposures from tobacco smoke: a review. *American journal of public health, 79*(8), 1044-1045. doi:10.2105/ajph.79.8.1044

Goniewicz, M. L., Knysak, J., Gawron, M., Kosmider, L., Sobczak, A., Kurek, J., . . . Benowitz, N. (2014). Levels of selected carcinogens and toxicants in vapour from electronic cigarettes. *Tob Control, 23*(2), 133-139. doi:10.1136/tobaccocontrol-2012-050859

Goniewicz, M. L., Smith, D. M., Edwards, K. C., Blount, B. C., Caldwell, K. L., Feng, J., . . . Hyland, A. J. (2018). Comparison of Nicotine and Toxicant Exposure in Users of Electronic Cigarettes and Combustible Cigarettes. *JAMA Netw Open, 1*(8), e185937. doi:10.1001/jamanetworkopen.2018.5937

Grana, R., Benowitz, N., & Glantz, S. A. (2014). E-cigarettes: a scientific review. *Circulation, 129*(19), 1972-1986. doi:10.1161/CIRCULATIONAHA.114.007667

Gravely, S., Driezen, P., Kyriakos, C. N., Thompson, M. E., Balmford, J., Demjen, T., . . . Consortium, E.-P. (2020). European adult smokers' perceptions of the

harmfulness of e-cigarettes relative to combustible cigarettes: cohort findings from the 2016 and 2018 EUREST-PLUS ITC Europe Surveys. *Eur J Public Health*. doi:10.1093/eurpub/ckz215

Green, M. J., Gray, L., Sweeting, H., & Benzeval, M. (2020). Socioeconomic patterning of vaping by smoking status among UK adults and youth. *BMC Public Health, 20*(1), 183. doi:10.1186/s12889-020-8270-3

Griffiths, A., Miller, J., Suzuki, D., Lewontin, R., & William, G. (2000). Mendel's experiments. In *An Introduction to Genetic Analysis* (7th Edition ed.). New York: W. H. Freeman.

Haapasalo, M. (2016). Level of evidence in endodontics: what does it mean? *Endodontic Topics, 34*(1), 30-41. doi:10.1111/etp.12099

Haidich, A. B. (2010). Meta-analysis in medical research. *Hippokratia, 14*(Suppl 1), 29-37. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3049418/

Hajek, P., Phillips-Waller, A., Przulj, D., Pesola, F., Smith, K. M., Bisal, N., . . . McRobbie, H. J. (2019). A Randomized Trial of E-Cigarettes versus Nicotine-Replacement Therapy. *New England Journal of Medicine, 380*(7), 629-637. doi:10.1056/NEJMoa1808779

Hallingberg, B., Maynard, O. M., Bauld, L., Brown, R., Gray, L., Lowthian, E., . . . Moore, G. (2020). Have e-cigarettes renormalised or displaced youth smoking? Results of a segmented regression analysis of repeated cross sectional survey data in England, Scotland and Wales. *Tobacco Control, 29*(2), 207-216. doi:10.1136/tobaccocontrol-2018-054584

Halpern-Felsher, B., Baker, M., & Stitzel, S. (2016). Decision-Making in Adolescents and Young Adults. In M. A. Diefenbach, S. Miller-Halegoua, & D. J. Bowen (Eds.), *Handbook of Health Decision Science* (pp. 157-167). New York, NY: Springer New York.

Hammond, D., Reid, J. L., Cole, A. G., & Leatherdale, S. T. (2017). Electronic cigarette use and smoking initiation among youth: a longitudinal cohort study. *Canadian Medical Association Journal, 189*(43), E1328-E1336. doi:10.1503/cmaj.161002

Hammond, D., Reid, J. L., Rynard, V. L., Fong, G. T., Cummings, K. M., McNeill, A., . . . White, C. M. (2019). Prevalence of vaping and smoking among adolescents in Canada, England, and the United States: repeat national cross sectional surveys. *BMJ, 365*, l2219. doi:10.1136/bmj.l2219

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform, 42*(2), 377-381. doi:10.1016/j.jbi.2008.08.010

Hartmann-Boyce, J., McRobbie, H., Lindson, N., Bullen, C., Begh, R., Theodoulou, A., . . . et al. (2020). Electronic cigarettes for smoking cessation. *Cochrane Database of Systematic Reviews*(10). doi:10.1002/14651858.CD010216.pub4

References

Hartmann-Boyce, J., McRobbie, H., Bullen, C., Begh, R., Stead, L. F., & Hajek, P. (2016). Electronic cigarettes for smoking cessation. *Cochrane Database of Systematic Reviews*(9). doi:10.1002/14651858.CD010216.pub3

Hartwell, G., Thomas, S., Egan, M., Gilmore, A., & Petticrew, M. (2017). E-cigarettes and equity: a systematic review of differences in awareness and use between sociodemographic groups. *Tobacco Control, 26*(E2), e85-e91. doi:10.1136/tobaccocontrol-2016-053222

Hartwig, F. P., Davey Smith, G., & Bowden, J. (2017). Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol, 46*(6), 1985-1998. doi:10.1093/ije/dyx102

Hatsukami, D., Pentel, P. R., Jensen, J., Nelson, D., Allen, S. S., Goldman, A., & Rafael, D. (1998). Cotinine: effects with and without nicotine. *Psychopharmacology (Berl), 135*(2), 141-150. doi:10.1007/s002130050495

Heath, A. C., Madden, P. A., Slutske, W. S., & Martin, N. G. (1995). Personality and the inheritance of smoking behavior: a genetic perspective. *Behav Genet, 25*(2), 103-117. doi:10.1007/bf02196921

Heatherton, T. F., Kozlowski, L. T., Frecker, R. C., & Fagerstrom, K. O. (1991). The Fagerstrom Test for Nicotine Dependence: a revision of the Fagerstrom Tolerance Questionnaire. *Br J Addict, 86*(9), 1119-1127. doi:10.1111/j.1360-0443.1991.tb01879.x

Hébert-Losier, A., Filion, K. B., Windle, S. B., & Eisenberg, M. J. (2020). A Randomized Controlled Trial Evaluating the Efficacy of E-Cigarette Use for Smoking Cessation in the General Population: E3 Trial Design. *CJC Open, 2*(3), 168-175. doi:https://doi.org/10.1016/j.cjco.2020.03.006

Hedman, L., Backman, H., Stridsman, C., Bosson, J. A., Lundback, M., Lindberg, A., . . . Ekerljung, L. (2018). Association of Electronic Cigarette Use With Smoking Habits, Demographic Factors, and Respiratory Symptoms. *Jama Network Open, 1*(3). doi:10.1001/jamanetworkopen.2018.0789

Heinzerling, A., Armatas, C., Karmarkar, E., Attfield, K., Guo, W., Wang, Y., . . . Wilken, J. A. (2020). Severe Lung Injury Associated With Use of e-Cigarette, or Vaping, Products—California, 2019. *JAMA Internal Medicine, 180*(6), 861-869. doi:10.1001/jamainternmed.2020.0664

Hemani, G., Bowden, J., & Davey Smith, G. (2018). Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum Mol Genet, 27*(R2), R195-R208. doi:10.1093/hmg/ddy163

Hemani, G., Bowden, J., Haycock, P., Zheng, J., Davis, O., Flach, P., . . . Davey Smith, G. (2017). Automating Mendelian randomization through machine learning to construct a putative causal map of the human phenome. *bioRxiv*, 173682. doi:10.1101/173682

References

Higgins, J. P. T., Green, S., & Cochrane Collaboration. (2008). *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley-Blackwell.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327*(7414), 557-560. doi:10.1136/bmj.327.7414.557

Hill, A. B. (1965). The Environment and Disease: Association or Causation? *Proc R Soc Med, 58*, 295-300. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/14283879

Hiscock, R., Bauld, L., Amos, A., Fidler, J. A., & Munafo, M. (2012). Socioeconomic status and smoking: a review. *Addiction Reviews, 1248*, 107-123. doi:10.1111/j.1749-6632.2011.06202.x

Hobbs, B. D., de Jong, K., Lamontagne, M., Bosse, Y., Shrine, N., Artigas, M. S., . . . International, C. G. C. (2017). Genetic loci associated with chronic obstructive pulmonary disease overlap with loci for lung function and pulmonary fibrosis. *Nat Genet, 49*(3), 426-432. doi:10.1038/ng.3752

Hoggart, C. J., Clark, T. G., De Iorio, M., Whittaker, J. C., & Balding, D. J. (2008). Genome-wide significance for dense SNP and resequencing data. *Genetic Epidemiology, 32*(2), 179-185. doi:10.1002/gepi.20292

House of Commons Science and Technology Committee. (2018). *E-cigarettes*. Retrieved from https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/505/505.pdf

Hughes, J. R., & Hatsukami, D. (1986). Signs and symptoms of tobacco withdrawal. *Arch Gen Psychiatry, 43*(3), 289-294. doi:10.1001/archpsyc.1986.01800030107013

Jackson, S. E., Hill, E., Shahab, L., Beard, E., Michie, S., & Brown, J. (2019). Prevalence and correlates of long-term e-cigarette and nicotine replacement therapy use: a prospective study in England. *BMJ Open, 9*(10), e029252. doi:10.1136/bmjopen-2019-029252

Jacobs, M. (2018). Adolescent smoking: The relationship between cigarette smoking, e-cigarette smoking and BMI. *Journal of Nutrition and Human Health, 2*(2). doi:10.35841/nutrition-human-health.2.2.17-23

Jakobsen, J. C., Gluud, C., Wetterslev, J., & Winkel, P. (2017). When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. *BMC Medical Research Methodology, 17*(1), 162. doi:10.1186/s12874-017-0442-1

Jamal, A., King, B. A., Neff, L. J., Whitmill, J., Babb, S. D., & Graffunder, C. M. (2016). Current Cigarette Smoking Among Adults - United States, 2005-2015. *Mmwr-Morbidity and Mortality Weekly Report, 65*(44), 1205-1211. doi:DOI 10.15585/mmwr.mm6544a2

References

Jayes, L., Haslam, P. L., Gratziou, C. G., Powell, P., Britton, J., Vardavas, C., . . . Tobacco Control Committee of the European Respiratory, S. (2016). SmokeHaz: Systematic Reviews and Meta-analyses of the Effects of Smoking on Respiratory Health. *Chest, 150*(1), 164-179. doi:10.1016/j.chest.2016.03.060

Jensen, R. P., Luo, W., Pankow, J. F., Strongin, R. M., & Peyton, D. H. (2015). Hidden formaldehyde in e-cigarette aerosols. *N Engl J Med, 372*(4), 392-394. doi:10.1056/NEJMc1413069

Jiang, Y., Chen, S., McGuire, D., Chen, F., Liu, M., Iacono, W. G., . . . Liu, D. J. (2018). Proper conditional analysis in the presence of missing data: Application to large scale meta-analysis of tobacco use phenotypes. *PLOS Genetics, 14*(7), e1007452. doi:10.1371/journal.pgen.1007452

Jo, Y.-H., Talmage, D. A., & Role, L. W. (2002). Nicotinic receptor-mediated effects on appetite and food intake. *Journal of neurobiology, 53*(4), 618-632. doi:10.1002/neu.10147

Jorenby, D. E., Hatsukami, D. K., Smith, S. S., Fiore, M. C., Allen, S., Jensen, J., & Baker, T. B. (1996). Characterization of tobacco withdrawal symptoms: transdermal nicotine reduces hunger and weight gain. *Psychopharmacology (Berl), 128*(2), 130-138. doi:10.1007/s002130050118

Kanwal, R., Kullman, G., Piacitelli, C., Boylstein, R., Sahakian, N., Martin, S., . . . Kreiss, K. (2006). Evaluation of Flavorings-Related Lung Disease Risk at Six Microwave Popcorn Plants. *Journal of Occupational and Environmental Medicine, 48*(2), 149-157. doi:10.1097/01.jom.0000194152.48728.fb

Karlsen, S., Millward, D., & Sandford, A. (2012). Investigating ethnic differences in current cigarette smoking over time using the health surveys for England. *Eur J Public Health, 22*(2), 254-256. doi:10.1093/eurpub/ckr097

Kennedy, C. D., van Schalkwyk, M. C. I., McKee, M., & Pisinger, C. (2019). The cardiovascular effects of electronic cigarettes: A systematic review of experimental studies. *Preventive Medicine, 127*, 105770. doi:https://doi.org/10.1016/j.ypmed.2019.105770

Khan, A., Shabbir, K., Ansari, J. K., & Zia, N. (2010). Comparison of forced expiratory volume in one second (FEV1) among asymptomatic smokers and non-smokers. *J Pak Med Assoc, 60*(3), 209-213. Retrieved from http://dx.doi.org/

Khouja, J. N., Munafò, M. R., Relton, C. L., Taylor, A. E., Gage, S. H., & Richmond, R. C. (2020). Investigating the added value of biomarkers compared with self-reported smoking in predicting future e-cigarette use: Evidence from a longitudinal UK cohort study. *PLOS ONE, 15*(7), e0235629. doi:10.1371/journal.pone.0235629

Khouja, J. N., Suddell, S. F., Peters, S. E., Taylor, A. E., & Munafo, M. R. (2020). Is e-cigarette use in non-smoking young adults associated with later smoking? A systematic review and meta-analysis. *Tob Control*. doi:10.1136/tobaccocontrol-2019-055433

References

Khouja, J. N., Taylor, A. E., & Munafò, M. R. (2020). Associations between reasons for vaping and current vaping and smoking status: Evidence from a UK based cohort. 19006007. doi:10.1101/19006007 %J medRxiv

Kim, H.-G., Cheon, E.-J., Bai, D.-S., Lee, Y. H., & Koo, B.-H. (2018). Stress and Heart Rate Variability: A Meta-Analysis and Review of the Literature. *Psychiatry investigation, 15*(3), 235-245. doi:10.30773/pi.2017.08.17

Kim, S., & Selya, A. S. (2019). The Relationship Between Electronic Cigarette Use and Conventional Cigarette Smoking Is Largely Attributable to Shared Risk Factors. *Nicotine & Tobacco Research*. doi:10.1093/ntr/ntz157

Kimber, C., Frings, D., Cox, S., Albery, I. P., & Dawkins, L. (2020). Communicating the relative health risks of E-cigarettes: An online experimental study exploring the effects of a comparative health message versus the EU nicotine addiction warnings on smokers' and non-smokers' risk perceptions and behavioural intentions. *Addictive Behaviors, 101*, 106177. doi:https://doi.org/10.1016/j.addbeh.2019.106177

Klein, D. E., Chaiton, M., Kundu, A., & Schwartz, R. (2020). A Literature Review on International E-cigarette Regulatory Policies. *Current Addiction Reports, 7*(4), 509-519. doi:10.1007/s40429-020-00332-w

Kong, G., Morean, M. E., Cavallo, D. A., Camenga, D. R., & Krishnan-Sarin, S. (2015). Reasons for Electronic Cigarette Experimentation and Discontinuation Among Adolescents and Young Adults. *Nicotine & Tobacco Research, 17*(7), 847-854. doi:10.1093/ntr/ntu257

Kreiss, K., Gomaa, A., Kullman, G., Fedan, K., Simoes, E. J., & Enright, P. L. (2002). Clinical Bronchiolitis Obliterans in Workers at a Microwave-Popcorn Plant. *New England Journal of Medicine, 347*(5), 330-338. doi:10.1056/NEJMoa020300

Kuchenbaecker, K. B., Hopper, J. L., Barnes, D. R., Phillips, K. A., Mooij, T. M., Roos-Blom, M. J., . . . Olsson, H. (2017). Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers. *Jama, 317*(23), 2402-2416. doi:10.1001/jama.2017.7112

Lai, S. H., Lai, H., Page, J. B., & McCoy, C. B. (2000). The association between cigarette smoking and drug abuse in the United States. *Journal of Addictive Diseases, 19*(4), 11-24. doi:DOI 10.1300/J069v19n04_02

Landry, R. L., Groom, A. L., Vu, T.-H. T., Stokes, A. C., Berry, K. M., Kesh, A., . . . Payne, T. J. (2019). The role of flavors in vaping initiation and satisfaction among U.S. adults. *Addictive Behaviors, 99*, 106077. doi:https://doi.org/10.1016/j.addbeh.2019.106077

Lawlor, D. A. (2016). Commentary: Two-sample Mendelian randomization: opportunities and challenges. *International journal of epidemiology, 45*(3), 908-915. doi:10.1093/ije/dyw127

References

Lawlor, D. A., Harbord, R. M., Sterne, J. A. C., Timpson, N., & Davey Smith, G. (2008). Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Statistics in Medicine, 27*(8), 1133-1163. doi:10.1002/sim.3034

Lawlor, D. A., Tilling, K., & Davey Smith, G. (2016). Triangulation in aetiological epidemiology. In *Int J Epidemiol* (Vol. 45, pp. 1866-1886).

Lawlor, D. A., Wade, K., Borges, M. C., Palmer, T., Hartwig, F. P., Hemani, G., & Bowden, J. (2019). A Mendelian Randomization dictionary: Useful definitions and descriptions for undertaking, understanding and interpreting Mendelian Randomization studies. doi: https://doi.org/10.31219/osf.io/6yzs7

Lederer, W., & Küchenhoff, H. (2006). A short Introduction to the SIMEX and MCSIMEX. *R News, 6*, 26-31.

Lee, J. A., Lee, S., & Cho, H. J. (2017). The Relation between Frequency of E-Cigarette Use and Frequency and Intensity of Cigarette Smoking among South Korean Adolescents. In *Int J Environ Res Public Health* (Vol. 14).

Lee, K. W. K., Richmond, R., Hu, P. Z., French, L., Shin, J., Bourdon, C., . . . Pausova, Z. (2015). Prenatal Exposure to Maternal Cigarette Smoking and DNA Methylation: Epigenome-Wide Association in a Discovery Sample of Adolescents and Replication in an Independent Cohort at Birth through 17 Years of Age. *Environmental Health Perspectives, 123*(2), 193-199. doi:10.1289/ehp.1408614

Leventhal, A. M., Strong, D. R., Kirkpatrick, M. G., Unger, J. B., Sussman, S., Riggs, N. R., . . . Audrain-McGovern, J. (2015). Association of Electronic Cigarette Use With Initiation of Combustible Tobacco Product Smoking in Early Adolescence. *Jama, 314*(7), 700-707. doi:10.1001/jama.2015.8950

Levy, D. T., Yuan, Z., & Li, Y. (2017). The Prevalence and Characteristics of E-Cigarette Users in the U.S. *Int J Environ Res Public Health, 14*(10). doi:10.3390/ijerph14101200

Lewis, C. M., & Vassos, E. (2020). Polygenic risk scores: from research tools to clinical instruments. *Genome Medicine, 12*(1), 44. doi:10.1186/s13073-020-00742-5

Linneberg, A., Jacobsen, R. K., Skaaby, T., Taylor, A. E., Fluharty, M. E., Jeppesen, J. L., . . . Husemoen, L. L. N. (2015). Effect of Smoking on Blood Pressure and Resting Heart Rate. *Circulation: Cardiovascular Genetics, 8*(6), 832-841. doi:doi:10.1161/CIRCGENETICS.115.001225

Lipsitch, M., Tchetgen Tchetgen, E., & Cohen, T. (2010). Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology (Cambridge, Mass.), 21*(3), 383-388. doi:10.1097/EDE.0b013e3181d61eeb

Liu, M. Z., Jiang, Y., Wedow, R., Li, Y., Brazel, D. M., Chen, F., . . . Psychiat, H. A. I. (2019). Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nature Genetics, 51*(2), 237-+. doi:10.1038/s41588-018-0307-5

References

Loukas, A., Marti, C. N., Cooper, M., Pasch, K. E., & Perry, C. L. (2018). Exclusive e-cigarette use predicts cigarette initiation among college students. *Addict Behav, 76*, 343-347. doi:10.1016/j.addbeh.2017.08.023

Lozano, P., Barrientos-Gutierrez, I., Arillo-Santillan, E., Morello, P., Mejia, R., Sargent, J. D., & Thrasher, J. F. (2017). A longitudinal study of electronic cigarette use and onset of conventional cigarette smoking and marijuana use among Mexican adolescents. *Drug and Alcohol Dependence, 180*, 427-430. doi:10.1016/j.drugalcdep.2017.09.001

MacDonald, A., & Middlekauff, H. R. (2019). Electronic cigarettes and cardiovascular health: what do we know so far? *Vasc Health Risk Manag, 15*, 159-174. doi:10.2147/VHRM.S175970

Marcon, A., Pesce, G., Calciano, L., Bellisario, V., Dharmage, S. C., Garcia-Aymerich, J., . . . Stud, A. L. E. C. (2018). Trends in smoking initiation in Europe over 40 years: A retrospective cohort study. *Plos One, 13*(8). doi:10.1371/journal.pone.0201881

Martens, E. P., Pestman, W. R., de Boer, A., Belitser, S. V., & Klungel, O. H. (2006). Instrumental variables: application and limitations. *Epidemiology, 17*(3), 260-267. doi:10.1097/01.ede.0000215160.88317.cb

McDonagh, M., Peterson, K., Raina, P., Chang, S., & Shekelle, P. (2013). Avoiding Bias in Selecting Studies. . In M. U. Oregon Health & Science University, and Southern California Evidence-based Practice Centers (Ed.), *Methods Guide for Comparative Effectiveness Reviews.*: AHRQ.

McNeill, A., Brose, L. S., Calder, R., Hitchman, S. C., Hajek, P., & McRobbie, H. (2015). *E-cigarettes: an evidence update*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/457102/Ecigarettes_an_evidence_update_A_report_commissioned_by_Public_Health_England_FINAL.pdf

Miech, R. A., Patrick, M. E., O'Malley, P. M., & Johnston, L. D. (2017). E-cigarette use as a predictor of cigarette smoking: results from a 1-year follow-up of a national sample of 12th grade students. doi:10.1136/tobaccocontrol-2016-053291

Millard, L. A. C., Munafò, M. R., Tilling, K., Wootton, R. E., & Smith, G. D. (2018). MR-pheWAS with stratification and interaction: Searching for the causal effects of smoking heaviness identified an effect on facial aging. *bioRxiv*, 441907. doi:10.1101/441907

Mineur, Y. S., Abizaid, A., Rao, Y., Salas, R., DiLeone, R. J., Gündisch, D., . . . Picciotto, M. R. (2011). Nicotine decreases food intake through activation of POMC neurons. *Science (New York, N.Y.), 332*(6035), 1330-1332. doi:10.1126/science.1201889

Minichino, A., Bersani, F. S., Calo, W. K., Spagnoli, F., Francesconi, M., Vicinanza, R., . . . Biondi, M. (2013). Smoking Behaviour and Mental Health Disorders-Mutual Influences and Implications for Therapy. *International Journal of Environmental Research and Public Health, 10*(10), 4790-4811. doi:10.3390/ijerph10104790

References

Mitchell, R., Hemani, G., Dudding, T., Corbin, L., Harrison, S., & Paternoster, L. (2019). UK Biobank Genetic Data: MRC-IEU Quality Control, version 2. doi:https://doi.org/10.5523/bris.1ovaau5sxunp2cv8rcy88688v

Mizes, J. S., Sloan, D. M., Segraves, K., Spring, B., Pingitore, R., & Kristeller, J. (1998). The influence of weight-related variables on smoking cessation. *Behavior Therapy, 29*(3), 371-385. doi:https://doi.org/10.1016/S0005-7894(98)80038-5

Moran, V. E. (2012). Cotinine: Beyond that Expected, More than a Biomarker of Tobacco Consumption. *Frontiers in pharmacology, 3*, 173-173. doi:10.3389/fphar.2012.00173

Morgenstern, M., Nies, A., Goecke, M., & Hanewinkel, R. (2018). E-Cigarettes and the Use of Conventional Cigarettes. *Dtsch Arztebl Int, 115*(14), 243-248. doi:10.3238/arztebl.2018.0243

Munafo, M. R., & Davey Smith, G. (2018). Robust research needs many lines of evidence. *Nature, 553*(7689), 399-401. doi:10.1038/d41586-018-01023-3

Munafò, M. R., Tilling, K., Taylor, A. E., Evans, D. M., & Davey Smith, G. (2018). Collider scope: when selection bias can substantially influence observed associations. *Int J Epidemiol, 47*(1), 226-235. doi:10.1093/ije/dyx206

Najem, B., Houssière, A., Pathak, A., Janssen, C., Lemogoum, D., Xhaët, O., . . . van de Borne, P. (2006). Acute Cardiovascular and Sympathetic Effects of Nicotine Replacement Therapy. *Hypertension, 47*(6), 1162-1167. doi:doi:10.1161/01.HYP.0000219284.47970.34

NHS. (2019a). Causes: Chronic obstructive pulmonary disease (COPD). Retrieved from https://www.nhs.uk/conditions/chronic-obstructive-pulmonary-disease-copd/causes/

NHS. (2019b). Stop smoking without putting on weight. Retrieved from https://www.nhs.uk/live-well/quit-smoking/stop-smoking-without-putting-on-weight/

Nicksic, N. E., Snell, L. M., & Barnes, A. J. (2019). Reasons to use e-cigarettes among adults and youth in the Population Assessment of Tobacco and Health (PATH) study. *Addictive Behaviors, 93*, 93-99. doi:10.1016/j.addbeh.2019.01.037

Nitsch, D., Molokhia, M., Smeeth, L., DeStavola, B. L., Whittaker, J. C., & Leon, D. A. (2006). Limits to causal inference based on Mendelian randomization: a comparison with randomized controlled trials. *Am J Epidemiol, 163*(5), 397-403. doi:10.1093/aje/kwj062

Nohmi, T. (2018). Thresholds of Genotoxic and Non-Genotoxic Carcinogens. *Toxicological research, 34*(4), 281-290. doi:10.5487/TR.2018.34.4.281

Northstone, K., Lewcock, M., Groom, A., Boyd, A., Macleod, J., Timpson, N., & Wells, N. (2019). The Avon Longitudinal Study of Parents and Children (ALSPAC): an

update on the enrolled sample of index children in 2019. *Wellcome Open Res, 4*, 51. doi:10.12688/wellcomeopenres.15132.1

Osei, A. D., Mirbolouk, M., Orimoloye, O. A., Dzaye, O., Uddin, S. M. I., Benjamin, E. J., . . . Blaha, M. J. (2020). Association Between E-Cigarette Use and Chronic Obstructive Pulmonary Disease by Smoking Status: Behavioral Risk Factor Surveillance System 2016 and 2017. *American Journal of Preventive Medicine, 58*(3), 336-342. doi:10.1016/j.amepre.2019.10.014

Panagiotou, O. A., Ioannidis, J. P. A., & the Genome-Wide Significance Project. (2011). What should the genome-wide significance threshold be? Empirical replication of borderline genetic associations. *International Journal of Epidemiology, 41*(1), 273-286. doi:10.1093/ije/dyr178

Paternoster, L., Tilling, K., & Davey Smith, G. (2017). Genetic epidemiology and Mendelian randomization for informing disease therapeutics: Conceptual and methodological challenges. *PLOS Genetics, 13*(10), e1006944. doi:10.1371/journal.pgen.1006944

Pearson, J. L., Richardson, A., Niaura, R. S., Vallone, D. M., & Abrams, D. B. (2012). e-Cigarette Awareness, Use, and Harm Perceptions in US Adults. *Am J Public Health, 102*(9), 1758-1766. doi:10.2105/ajph.2011.300526

Perialathan, K., Rahman, A. B., Lim, K. H., Adon, Y., Ahmad, A., Juatan, N., & Jaafar, N. (2018). Prevalence and associated factors of ever use of electronic cigarettes: Findings from a hospitals and health clinics study based in Malaysia. *Tobacco Induced Diseases, 16*(November). doi:10.18332/tid/99258

Perkins, K. A., Epstein, L. H., Sexton, J. E., Solberg-Kassel, R., Stiller, R. L., & Jacob, R. G. (1992). Effects of nicotine on hunger and eating in male and female smokers. *Psychopharmacology (Berl), 106*(1), 53-59. doi:10.1007/bf02253588

Perkins, K. A., Epstein, L. H., Stiller, R. L., Fernstrom, M. H., Sexton, J. E., Jacob, R. G., & Solberg, R. (1991). Acute effects of nicotine on hunger and caloric intake in smokers and nonsmokers. *Psychopharmacology (Berl), 103*(1), 103-109. doi:10.1007/bf02244083

Pierce, J. S., Abelmann, A., & Finley, B. L. (2014). Response to Muge Akpinar-Elci Letter to Editor re. Pierce et al., Diacetyl and 2,3-pentanedione exposures associated with cigarette smoking: implications for risk assessment of food and flavoring workers, Crit. Rev. Toxicol. 44(5):420-435, 2014. *Crit Rev Toxicol, 44*(7), 640-641. doi:10.3109/10408444.2014.930093

Pierce, J. S., Abelmann, A., Spicer, L. J., Adams, R. E., & Finley, B. L. (2014). Diacetyl and 2,3-pentanedione exposures associated with cigarette smoking: implications for risk assessment of food and flavoring workers. *Crit Rev Toxicol, 44*(5), 420-435. doi:10.3109/10408444.2014.882292

Pilhatsch, M., Scheuing, H., Kroemer, N., Kobiella, A., Bidlingmaier, M., Farger, G., . . . Zimmermann, U. S. (2014). Nicotine administration in healthy non-smokers

reduces appetite but does not alter plasma ghrelin. *Hum Psychopharmacol, 29*(4), 384-387. doi:10.1002/hup.2405

Polfus, L. M., Smith, J. A., Shimmin, L. C., Bielak, L. F., Morrison, A. C., Kardia, S. L., . . . Hixson, J. E. (2013). Genome-wide association study of gene by smoking interactions in coronary artery calcification. *PLoS One, 8*(10), e74642. doi:10.1371/journal.pone.0074642

Polosa, R., Cibella, F., Caponnetto, P., Maglia, M., Prosperini, U., Russo, C., & Tashkin, D. (2017). Health impact of E-cigarettes: a prospective 3.5-year study of regular daily users who have never smoked. *Sci Rep, 7*(1), 13825. doi:10.1038/s41598-017-14043-2

Primack, B. A., Shensa, A., Sidani, J. E., Hoffman, B. L., Soneji, S., Sargent, J. D., . . . Fine, M. J. (2018). Initiation of Traditional Cigarette Smoking after Electronic Cigarette Use Among Tobacco-Naive US Young Adults. *Am J Med, 131*(4), 443.e441-443.e449. doi:10.1016/j.amjmed.2017.11.005

Primack, B. A., Shenza, A., Sidani, J. E., Hoffman, B. L., Soneji, S., Fine, M. J., . . . Sargent, J. D. (2016). *Initiation of cigarette smoking after e-cigarette use: a nationally representative study*. Paper presented at the Socieety of Behavioural Medicine.

Primack, B. A., Soneji, S., Stoolmiller, M., Fine, M. J., & Sargent, J. D. (2015). Progression to Traditional Cigarette Smoking After Electronic Cigarette Use Among US Adolescents and Young Adults. *Jama Pediatrics, 169*(11), 1018-1023. doi:10.1001/jamapediatrics.2015.1742

Public Health England. (2015). *E-cigarettes: an evidence update*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/457102/Ecigarettes_an_evidence_update_A_report_commissioned_by_Public_Health_England_FINAL.pdf

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., . . . Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet, 81*(3), 559-575. doi:10.1086/519795

Qasim, H., Karim, Z. A., Rivera, J. O., Khasawneh, F. T., & Alshbool, F. Z. (2017). Impact of Electronic Cigarettes on the Cardiovascular System. *Journal of the American Heart Association, 6*(9), e006353. doi:doi:10.1161/JAHA.117.006353

Ramprasad, V., Santhosh, K., & Lim, A. (2018). Carbon monoxide breath analyzers and its role in tobacco cessation: A narrative review of literature. *Journal of International Oral Health, 10*(2), 71-76. doi:10.4103/jioh.jioh_273_17

Reed, E., Nunez, S., Kulp, D., Qian, J., Reilly, M. P., & Foulkes, A. S. (2015). A guide to genome-wide association analysis and post-analytic interrogation. *Statistics in Medicine, 34*(28), 3769-3792. doi:10.1002/sim.6605

Retraction to: Electronic Cigarette Use and Myocardial Infarction Among Adults in the US Population Assessment of Tobacco and Health. (2020). *J Am Heart Assoc, 9*(4), e014519. doi:10.1161/jaha.119.014519

References

Reyna, V. F., & Farley, F. (2006). Risk and rationality in adolescent decision making - Implications for theory, practice, and public policy. *Psychological Science*, 1-44. doi:Doi 10.1145/1142680.1142682

Rhoades, D. A., Comiford, A. L., Dvorak, J. D., Ding, K., Hopkins, M., Spicer, P., . . . Doescher, M. P. (2019). Vaping patterns, nicotine dependence and reasons for vaping among American Indian dual users of cigarettes and electronic cigarettes. *BMC Public Health, 19*(1), 1211. doi:10.1186/s12889-019-7523-5

Richmond, R., C., S. R., Khouja, J., Prince, C., Board, A., Sharp, G., . . . Gage, S. (2018). *Investigating DNA Methylation Signatures of E-cigarette Use*. Paper presented at the Society for Research on Nicotine and Tobacco, San Francisco. https://cdn.ymaws.com/www.srnt.org/resource/resmgr/SRNT19_Rapid_Abstracts.pdf

Ritchie, H., & Roser, M. (2019). *Causes of Death*. Retrieved from https://ourworldindata.org/causes-of-death

Rodu, B., & Plurphanswat, N. (2020). A re-analysis of e-cigarette use and heart attacks in PATH wave 1 data. *Addiction, n/a*(n/a). doi:10.1111/add.15067

Russo, C., Cibella, F., Caponnetto, P., Campagna, D., Maglia, M., Frazzetto, E., . . . Polosa, R. (2016). Evaluation of Post Cessation Weight Gain in a 1-Year Randomized Smoking Cessation Trial of Electronic Cigarettes. *Sci Rep, 6*, 18763. doi:10.1038/srep18763

Russo, C., Cibella, F., Mondati, E., Caponnetto, P., Frazzetto, E., Caruso, M., . . . Polosa, R. (2018). Lack of Substantial Post-Cessation Weight Increase in Electronic Cigarettes Users. *Int J Environ Res Public Health, 15*(4). doi:10.3390/ijerph15040581

Saddleson, M. L., Kozlowski, L. T., Giovino, G. A., Goniewicz, M. L., Mahoney, M. C., Homish, G. G., & Arora, A. (2016). Enjoyment and other reasons for electronic cigarette use: Results from college students in New York. *Addictive Behaviors, 54*, 33-39. doi:10.1016/j.addbeh.2015.11.012

Sæbø, G., & Scheffels, J. (2017). Assessing notions of denormalization and renormalization of smoking in light of e-cigarette regulation. *International Journal of Drug Policy, 49*, 58-64. doi:https://doi.org/10.1016/j.drugpo.2017.07.026

Sagone, E., & Caroli, M. E. D. (2013). The Influence of Creative Personality Factors on Interpersonal Adjustment in Adolescents: What's the Relationship? *Procedia - Social and Behavioral Sciences, 82*, 131-136. doi:https://doi.org/10.1016/j.sbspro.2013.06.236

Salamanca, J. C., Meehan-Atrash, J., Vreeke, S., Escobedo, J. O., Peyton, D. H., & Strongin, R. M. (2018). E-cigarettes can emit formaldehyde at high levels under conditions that have been reported to be non-averse to users. *Scientific reports, 8*(1), 7559-7559. doi:10.1038/s41598-018-25907-6

References

Samek, D. R., & Hicks, B. M. (2014). Externalizing Disorders and Environmental Risk: Mechanisms of Gene-Environment Interplay and Strategies for Intervention. *Clinical practice (London, England), 11*(5), 537-547. doi:10.2217/CPR.14.47

Sanderson, E., Davey Smith, G., Windmeijer, F., & Bowden, J. (2019). An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. *Int J Epidemiol, 48*(3), 713-727. doi:10.1093/ije/dyy262

Sanderson, E., Spiller, W., & Bowden, J. (2020). Testing and Correcting for Weak and Pleiotropic Instruments in Two-Sample Multivariable Mendelian Randomisation. *bioRxiv*, 2020.2004.2002.021980. doi:10.1101/2020.04.02.021980

Shahab, L., Dobbie, F., Hiscock, R., McNeill, A., & Bauld, L. (2017). Prevalence and Impact of Long-term Use of Nicotine Replacement Therapy in UK Stop-Smoking Services: Findings From the ELONS Study. *Nicotine Tob Res, 20*(1), 81-88. doi:10.1093/ntr/ntw258

Shahab, L., Goniewicz, M. L., Blount, B. C., Brown, J., McNeill, A., Alwis, K. U., . . . West, R. (2017). Nicotine, Carcinogen, and Toxin Exposure in Long-Term E-Cigarette and Nicotine Replacement Therapy Users: A Cross-sectional Study. *Annals of internal medicine, 166*(6), 390-400. doi:10.7326/M16-1107

Shahab, L., Jarvis, M. J., Britton, J., & West, R. (2006). Prevalence, diagnosis and relation to tobacco dependence of chronic obstructive pulmonary disease in a nationally representative population sample. *Thorax, 61*(12), 1043-1047. doi:10.1136/thx.2006.064410

Sharpe, D. (1997). Of apples and oranges, file drawers and garbage: Why validity issues in meta-analysis will not go away. *Clinical Psychology Review, 17*(8), 881-901. doi:https://doi.org/10.1016/S0272-7358(97)00056-1

Shi, H., Kichaev, G., & Pasaniuc, B. (2016). Contrasting the Genetic Architecture of 30 Complex Traits from Summary Association Data. *Am J Hum Genet, 99*(1), 139-153. doi:10.1016/j.ajhg.2016.05.013

Soneji, S., Barrington-Trimis, J. L., Wills, T. A., Levanthal, A. M., Unger, J. B., Gibson, L. A., . . . Sargent, J. D. (2017). Association between initial use of e-cigarettes and subsequent cigarette smoking among adolescents and young adults: A systematic review and meta-analysis. *JAMA Pediatrics*. doi:10.1001/jamapediatrics.2017.1488

Soneji, S., Barrington-Trimis, J. L., Wills, T. A., Leventhal, A. M., Unger, J. B., Gibson, L. A., . . . Sargent, J. D. (2017). Association Between Initial Use of e-Cigarettes and Subsequent Cigarette Smoking Among Adolescents and Young Adults: A Systematic Review and Meta-analysis. *JAMA Pediatr, 171*(8), 788-797. doi:10.1001/jamapediatrics.2017.1488

Spindle, T. R., Hiler, M. M., Cooke, M. E., Eissenberg, T., Kendler, K. S., & Dick, D. M. (2017). Electronic cigarette use and uptake of cigarette smoking: A longitudinal examination of U.S. college students. *Addict Behav, 67*, 66-72. doi:10.1016/j.addbeh.2016.12.009

References

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Lowe, B. (2006). A brief measure for assessing generalized anxiety disorder - The GAD-7. *Archives of Internal Medicine, 166*(10), 1092-1097. doi:DOI 10.1001/archinte.166.10.1092

Srivastava, E. D., Russell, M. A., Feyerabend, C., Masterson, J. G., & Rhodes, J. (1991). Sensitivity and tolerance to nicotine in smokers and nonsmokers. *Psychopharmacology (Berl), 105*(1), 63-68. doi:10.1007/BF02316865

Stanwick, R. (2015). E-cigarettes: Are we renormalizing public smoking? Reversing five decades of tobacco control and revitalizing nicotine dependency in children and youth in Canada. *Paediatrics & child health, 20*(2), 101-105. doi:10.1093/pch/20.2.101

StataCorp. (2017). Stata Statistical Software (Version 15.1). College Station, TX: StataCorp LLC.

Sterne, J. A., & Harbord, R. M. (2004). Funnel plots in meta-analysis. *Stata Journal, 4*(2), 127-141.

Sterne, J. A., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., . . . Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ, 338*, b2393. doi:10.1136/bmj.b2393

Swanson, S. A., Tiemeier, H., Ikram, M. A., & Hernán, M. A. (2017). Nature as a Trialist?: Deconstructing the Analogy Between Mendelian Randomization and Randomized Trials. *Epidemiology, 28*(5), 653-659. doi:10.1097/ede.0000000000000699

Tattan-Birch, H., Brown, J., Shahab, L., & Jackson, S. E. (2020). Association of the US Outbreak of Vaping-Associated Lung Injury With Perceived Harm of e-Cigarettes Compared With Cigarettes. *JAMA Network Open, 3*(6), e206981-e206981. doi:10.1001/jamanetworkopen.2020.6981

Taylor, A. E., Davies, N. M., Ware, J. J., VanderWeele, T., Smith, G. D., & Munafo, M. R. (2014). Mendelian randomization in health research: using appropriate genetic variants and avoiding biased estimates. *Econ Hum Biol, 13*, 99-106. doi:10.1016/j.ehb.2013.12.002

Taylor, A. E., Jones, H. J., Sallis, H., Euesden, J., Stergiakouli, E., Davies, N. M., . . . Tilling, K. (2018). Exploring the association of genetic factors with participation in the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology, 47*(4), 1207-1216. doi:10.1093/ije/dyy060 %J International Journal of Epidemiology

Taylor, A. E., Morris, R. W., Fluharty, M. E., Bjorngaard, J. H., Asvold, B. O., Gabrielsen, M. E., . . . Sattar, N. (2014). Stratification by smoking status reveals an association of CHRNA5-A3-B4 genotype with body mass index in never smokers. *PLoS Genet, 10*(12), e1004799. doi:10.1371/journal.pgen.1004799

References

Taylor, A. E., & Munafò, M. R. (2014). Commentary: Does mortality from smoking have implications for future Mendelian randomization studies? *Int J Epidemiol, 43*(5), 1483-1486. doi:10.1093/ije/dyu151

Taylor, A. E., Richmond, R. C., Palviainen, T., Loukola, A., Wootton, R. E., Kaprio, J., . . . Munafo, M. R. (2019). The effect of body mass index on smoking behaviour and nicotine metabolism: a Mendelian randomization study. *Hum Mol Genet, 28*(8), 1322-1330. doi:10.1093/hmg/ddy434

Thapar, A., Rice, F., Hay, D., Boivin, J., Langley, K., van den Bree, M., . . . Harold, G. (2009). Prenatal smoking might not cause attention-deficit/hyperactivity disorder: evidence from a novel design. *Biological psychiatry, 66*(8), 722-727. doi:10.1016/j.biopsych.2009.05.032

Tobacco and Genetics Consortium. (2010). Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet, 42*(5), 441-447. doi:10.1038/ng.571

Tobacco Products Directive. (2014). Retrieved from https://ec.europa.eu/health/tobacco/products/revision_en

Treur, J. L., Demontis, D., Smith, G. D., Sallis, H., Richardson, T. G., Wiers, R. W., . . . Munafo, M. R. (2019). Investigating causality between liability to ADHD and substance use, and liability to substance use and ADHD risk, using Mendelian randomization. *Addiction Biology*. doi:10.1111/adb.12849

Treur, J. L., Rozema, A. D., Mathijssen, J. J. P., van Oers, H., & Vink, J. M. (2018). E-cigarette and waterpipe use in two adolescent cohorts: cross-sectional and longitudinal associations with conventional cigarette smoking. *Eur J Epidemiol, 33*(3), 323-334. doi:10.1007/s10654-017-0345-9

Tsai, J., Walton, K., Coleman, B. N., Sharapova, S. R., Johnson, S. E., Kennedy, S. M., & Caraballo, R. S. (2018). Reasons for Electronic Cigarette Use Among Middle and High School Students - National Youth Tobacco Survey, United States, 2016. *Mmwr-Morbidity and Mortality Weekly Report, 67*(6), 196-200. Retrieved from <Go to ISI>://WOS:000425099000005

Vanyukov, M. M., Tarter, R. E., Kirillova, G. P., Kirisci, L., Reynolds, M. D., Kreek, M. J., . . . Ridenour, T. A. (2012). Common liability to addiction and "gateway hypothesis": Theoretical, empirical and evolutionary perspective. *Drug and Alcohol Dependence, 123*, S3-S17. doi:10.1016/j.drugalcdep.2011.12.018

Vie, G. Å., Wootton, R. E., Bjørngaard, J. H., Åsvold, B. O., Taylor, A. E., Gabrielsen, M. E., . . . Munafò, M. R. (2019). The effect of smoking intensity on all-cause and cause-specific mortality—a Mendelian randomization analysis. *International Journal of Epidemiology, 48*(5), 1438-1446. doi:10.1093/ije/dyz081

Ware, J. J., Chen, X., Vink, J., Loukola, A., Minica, C., Pool, R., . . . Munafo, M. R. (2016). Genome-Wide Meta-Analysis of Cotinine Levels in Cigarette Smokers Identifies Locus at 4q13.2. *Sci Rep, 6*, 20092. doi:10.1038/srep20092

References

Ware, J. J., van den Bree, M. B. M., & Munafò, M. R. (2011). Association of the CHRNA5-A3-B4 Gene Cluster With Heaviness of Smoking: A Meta-Analysis. *Nicotine & Tobacco Research, 13*(12), 1167-1175. doi:10.1093/ntr/ntr118

Watkins, S. L., Glantz, S. A., & Chaffee, B. W. (2018). Association of Noncigarette Tobacco Product Use With Future Cigarette Smoking Among Youth in the Population Assessment of Tobacco and Health (PATH) Study, 2013-2015. *JAMA Pediatr, 172*(2), 181-187. doi:10.1001/jamapediatrics.2017.4173

Wells, G. A., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., & Tugwell, P. (2012). The Newcastle-Ottawa Scale (NOS) for assessing the quality if nonrandomized studies in meta-analyses. Retrieved from http://wwwohrica/programs/clinical_epidemiology/oxfordasp

West, R. (2017). Tobacco smoking: Health impact, prevalence, correlates and interventions. *Psychol Health, 32*(8), 1018-1036. doi:10.1080/08870446.2017.1325890

West, R., Baker, C. L., Cappelleri, J. C., & Bushmakin, A. G. (2008). Effect of varenicline and bupropion SR on craving, nicotine withdrawal symptoms, and rewarding effects of smoking during a quit attempt. *Psychopharmacology (Berl), 197*(3), 371-377. doi:10.1007/s00213-007-1041-3

Williamson, D. F., Madans, J., Anda, R. F., Kleinman, J. C., Giovino, G. A., & Byers, T. (1991). Smoking cessation and severity of weight gain in a national cohort. *N Engl J Med, 324*(11), 739-745. doi:10.1056/NEJM199103143241106

Wills, T. A., Knight, R., Sargent, J. D., Gibbons, F. X., Pagano, I., & Williams, R. J. (2017). Longitudinal study of e-cigarette use and onset of cigarette smoking among high school students in Hawaii. *Tob Control, 26*(1), 34-39. doi:10.1136/tobaccocontrol-2015-052705

Wurtz, P., Wang, Q., Kangas, A. J., Richmond, R. C., Skarp, J., Tiainen, M., . . . Ala-Korpela, M. (2014). Metabolic signatures of adiposity in young adults: Mendelian randomization analysis and effects of weight change. *PLoS Med, 11*(12), e1001765. doi:10.1371/journal.pmed.1001765

Xie, Z., Ossip, D. J., Rahman, I., & Li, D. (2019). Use of Electronic Cigarettes and Self-Reported Chronic Obstructive Pulmonary Disease Diagnosis in Adults. *Nicotine & Tobacco Research, 22*(7), 1155-1161. doi:10.1093/ntr/ntz234

Yong, H. H., Borland, R., Cummings, K. M., Gravely, S., Thrasher, J. F., McNeill, A., . . . Fong, G. T. (2019). Reasons for regular vaping and for its discontinuation among smokers and recent ex-smokers: findings from the 2016 ITC Four Country Smoking and Vaping Survey. *Addiction, 114*, 35-48. doi:10.1111/add.14593

Zeilinger, S., Kuhnel, B., Klopp, N., Baurecht, H., Kleinschmidt, A., Gieger, C., . . . Illig, T. (2013). Tobacco Smoking Leads to Extensive Genome-Wide Changes in DNA Methylation. *Plos One, 8*(5). doi:10.1371/journal.pone.0063812

References

Zhang, R., Chu, M., Zhao, Y., Wu, C., Guo, H., Shi, Y., . . . Chen, F. (2014). A genome-wide gene-environment interaction analysis for tobacco smoke and lung cancer susceptibility. *Carcinogenesis, 35*(7), 1528-1535. doi:10.1093/carcin/bgu076

Zhao, H., Mitra, N., Kanetsky, P. A., Nathanson, K. L., & Rebbeck, T. R. (2018). A practical approach to adjusting for population stratification in genome-wide association studies: principal components and propensity scores (PCAPS). *Statistical applications in genetics and molecular biology, 17*(6), /j/sagmb.2018.2017.issue-2016/sagmb-2017-0054/sagmb-2017-0054.xml. doi:10.1515/sagmb-2017-0054

Zhu, S. H., Zhuang, Y. L., Wong, S. S., Cummins, S. E., & Tedeschi, G. J. (2017). E-cigarette use and associated changes in population smoking cessation: evidence from US current population surveys. *Bmj-British Medical Journal, 358*. doi:10.1136/bmj.j3262

List of Abbreviations

Abbreviation	Description
ACR	Assumed control risk
ADHD	Attention deficit hyperactivity disorder
ALSPAC	Avon longitudinal study of parents and children
ANDS	Alternative nicotine delivery system
ANOVA	Analysis of variance
ASH	Action on Smoking and Health
AUDIT	Alcohol use disorders identification test
BMI	Body mass index
BRCA1	Breast cancer type 1 gene
BRCA2	Breast cancer type 2 gene
CD	Conduct disorder
CDC	Centre for disease control and prevention
CHD	Coronary heart disease
CHRNA	Cholinergic receptor
CI	Confidence interval
CIS-r	Computerised interview schedule - revised
CO	Carbon monoxide
COPD	Chronic obstructive pulmonary disease
COVID-19	Coronavirus disease
CPD	Cigarettes per day
DAWBA	Development and wellbeing assessment
DEBQ	Dutch eating behaviour questionnaire
DNA	Deoxyribonucleic acid
DSM	Diagnostic and statistical manual of mental disorders
ENDS	Electronic nicotine delivery system
EU	European union
EVALI	E-cigarette and vaping associated lung injury
FEV-1	Forced expiratory volume in 1 second
FTND	Fagerström test of nicotine dependence
FVC	Forced vital capacity
GAD	Generalised anxiety disorder
GN	Genetic variants as proxies for nicotine
GRADE	Grading of recommendations assessment, development and evaluation
GS	Genetic variants as proxies for smoking heaviness
GSCAN	Genome-wide association study & sequencing consortium of alcohol and nicotine use
GSN	Genetic variants as proxies for nicotine and smoking heaviness
GWAS	Genome-wide association study
GX	Genetic variants as proxies for the exposure
HR	Heart rate
ICD	International statistical classification of diseases and related health problems
ID	Identification
IEU	Integrative epidemiology unit

List of Abbreviations

ITC	International Tobacco Control
IVW	Inverse variance weighted
LD	Linkage disequilibrium
LSD	Lysergic acid diethylamide
MAF	Minor allele frequency
MeSH	Medical subject headings
MHRA	Medicines and healthcare products regulatory agency
MNWS	Minnesota nicotine withdrawal scale
MOOSE	Meta-analyses of observational studies in epidemiology
MR	Mendelian randomisation
MRC	Medical research council
MVMR	Multivariable Mendelian randomisation
NHS	National health service
NIH	National institute of health
NOS	Newcastle-Ottawa scale
NRT	Nicotine replacement therapy
ODD	Oppositional defiant disorder
OPCS	Office of population censuses and surveys
OR	Odds ratio
OSF	Open science framework
POMC	Pro-opiomelanocortin gene
PRISMA	Preferred reporting items for systematic reviews and meta-analyses
PRS	Polygenic risk score
QC	Quality check
RCT	Randomised controlled trial
RDA	Rebuildable dripping atomiser
REC	Research ethics committee
RR	Risk ratio
RRR	Relative risk ratio
SD	Standard deviation
SDQ	Strengths and difficulties questionnaire
SE	Standard error
SEP	Socioeconomic position
SES	Socioeconomic status
SIMEX	Simulation extrapolation
SNP	Single nucleotide polymorphism
TAG	Tobacco and alcohol genetics consortium
TARG	Tobacco and alcohol research group
THC	Tetrahydrocannabinol
UK	United Kingdom
US	United States
USA	United States of America
USB	Universal serial bus
VAS	Visual analogue scale
YP	Young person

Appendices

Appendix 1. Systematic search strategy.

Search dates

Search 1 completed: 12/02/18

Search 2 completed: 09/11/18

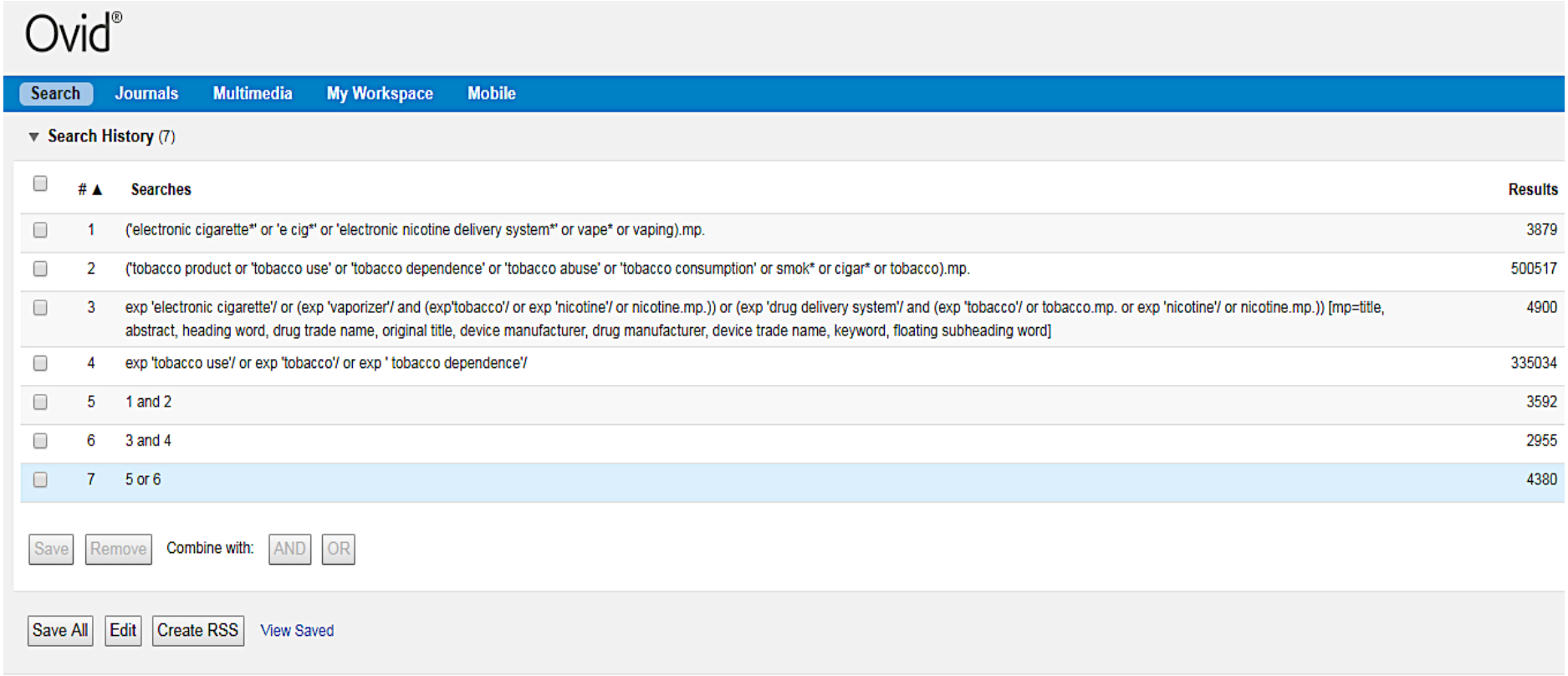
Search 3 completed: 24/11/18

Pubmed search terms

(Tobacco Use[mesh] OR Tobacco[mesh] OR Tobacco use disorder[mesh] OR Tobacco Products[mesh] OR Cigar*[tiab] OR Tobacco[tiab] OR Smok*[tiab]) AND (Electronic Cigarettes[mesh] OR (Nebulizers and Vaporizers[mesh] AND (Tobacco[mesh] OR Tobacco[tiab] OR Nicotine[mesh] OR Nicotine[tiab])) OR Electronic Cigarette*[tiab] OR E-Cig*[tiab] OR Electronic Nicotine Delivery System*[tiab] OR Vape[tiab] OR Vaping[tiab] OR Alternative Nicotine Delivery System*[tiab].

Appendices

Embase search terms



Ovid®

Search | Journals | Multimedia | My Workspace | Mobile

▼ Search History (7)

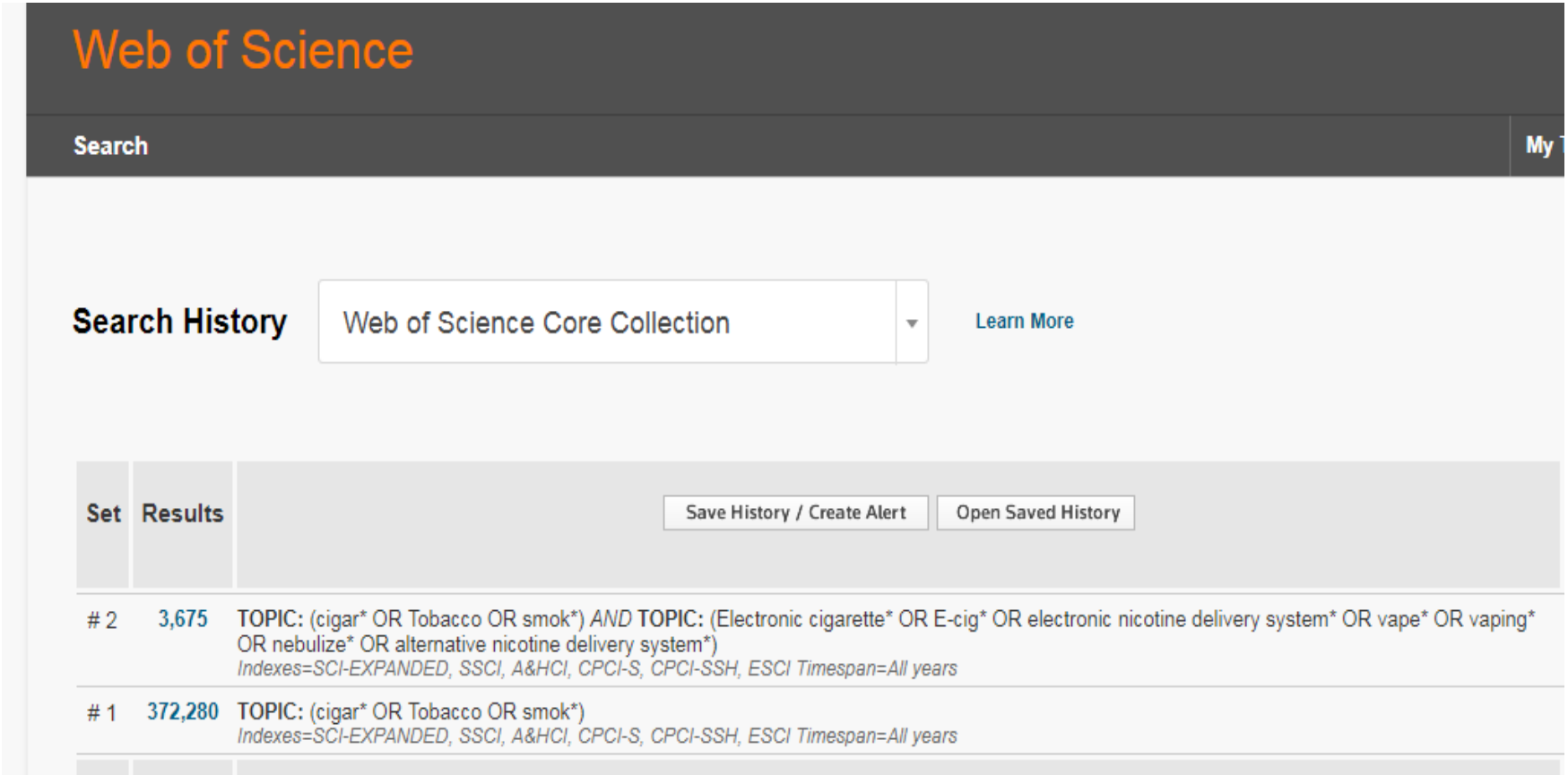
	# ▲	Searches	Results
☐	1	('electronic cigarette*' or 'e cig*' or 'electronic nicotine delivery system*' or vape* or vaping).mp.	3879
☐	2	('tobacco product or 'tobacco use' or 'tobacco dependence' or 'tobacco abuse' or 'tobacco consumption' or smok* or cigar* or tobacco).mp.	500517
☐	3	exp 'electronic cigarette'/ or (exp 'vaporizer'/ and (exp'tobacco'/ or exp 'nicotine'/ or nicotine.mp.)) or (exp 'drug delivery system'/ and (exp 'tobacco'/ or tobacco.mp. or exp 'nicotine'/ or nicotine.mp.)) [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word]	4900
☐	4	exp 'tobacco use'/ or exp 'tobacco'/ or exp ' tobacco dependence'/	335034
☐	5	1 and 2	3592
☐	6	3 and 4	2955
☐	7	5 or 6	4380

Save | Remove | Combine with: AND | OR

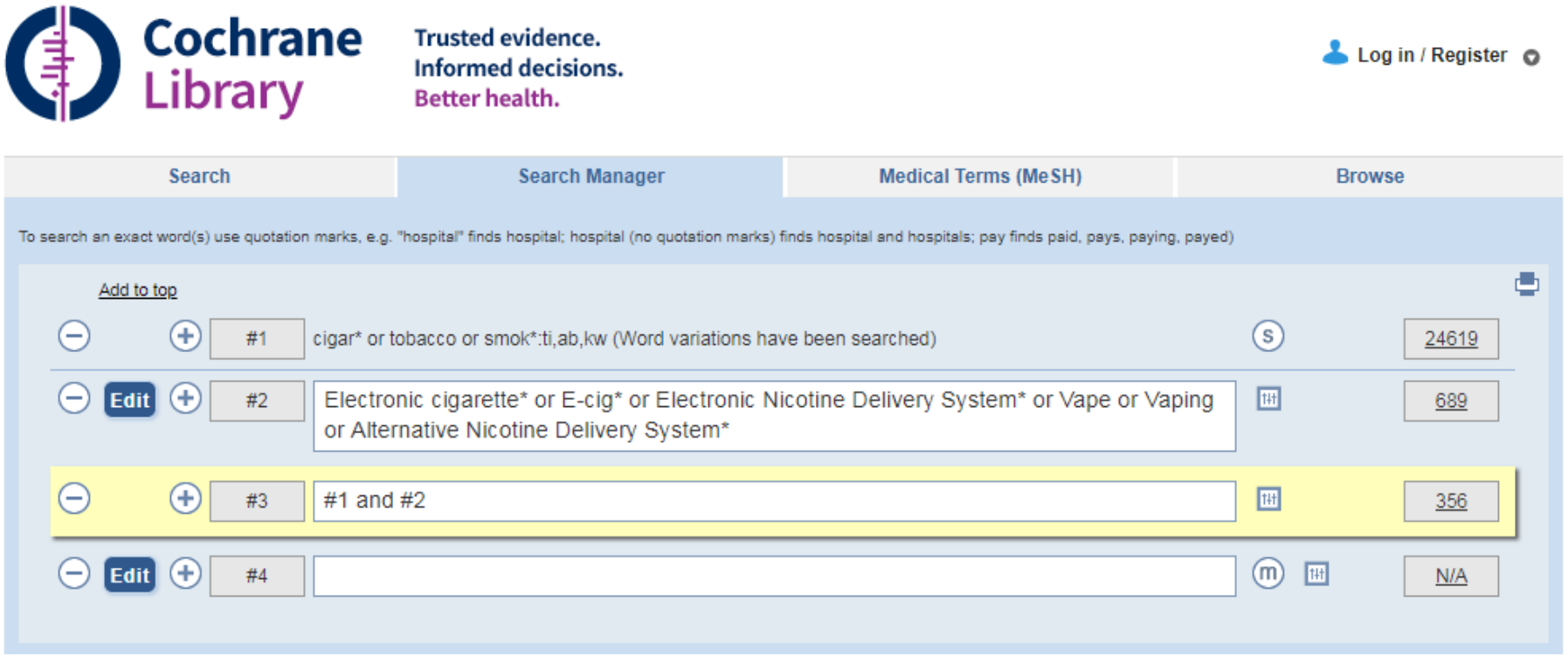
Save All | Edit | Create RSS | View Saved

Appendices

Web of Science search terms



Wiley Cochrance Library search terms



Appendices

Society for Research on Nicotine and Tobacco and the Society for Behavioural Medicine search terms

Tobacco

Smok*

Cigarette

E-cigarette

E-cig

Electronic Nicotine Delivery Device

Vape

Vaping

Appendix 2. Newcastle-Ottawa Quality Assessment Form for Cohort Studies.

Note: A study can be given a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability.

Selection
1) Representativeness of the exposed cohort
a) Truly representative of the average non-smoker who has used an e-cigarette in the community ***(one star)***
b) Somewhat representative of the average non-smoker who has used an e-cigarette in the community ***(one star)***
c) Selected group e.g., nurses, volunteers
d) No description of the derivation of the cohort

2) Selection of the non-exposed cohort
a) Drawn from the same community as the exposed cohort ***(one star)***
b) Drawn from a different source
c) No description of the derivation of the non-exposed cohort

3) Ascertainment of exposure
a) Secure record (e.g., stop smoking service record) ***(one star)***
b) Structured interview ***(one star)***
c) Written self-report
d) No description
e) Other

4) Demonstration that outcome of interest was not present at start of study
a) Yes ***(one star)***
b) No

Comparability
1) Comparability of cohorts on the basis of the design or analysis controlled for confounders
a) The study controls for age ***(one star)***
b) Study controls for additional factors (e.g., sex and SES) ***(one star)***
c) Study controls for A and B ***(two stars)***
d) Cohorts are not comparable on the basis of the design or analysis controlled for confounders

Outcome
1) Assessment of outcome
a) Independent blind assessment ***(one star)***
b) Record linkage ***(one star)***
c) Self-report
d) No description
e) Other

2) Was follow-up long enough for outcomes to occur
a) Yes (minimum 3 months) ***(one star)***
b) No

3) Adequacy of follow-up of cohorts
a) Complete follow up- all subject accounted for ***(one star)***
b) Subjects lost to follow up unlikely to introduce bias- number lost less than or equal to 30% or description of those lost suggested no different from those followed ***(one star)***
c) Follow up rate less than 70% and no description of those lost
d) No statement

Thresholds for converting the Newcastle-Ottawa scales to AHRQ standards (good, fair, and poor):

Good quality: 3 or 4 stars in selection domain AND 1 or 2 stars in comparability domain AND 2 or 3 stars in outcome/exposure domain

Fair quality: 2 stars in selection domain AND 1 or 2 stars in comparability domain AND 2 or 3 stars in outcome/exposure domain

Poor quality: 0 or 1 star in selection domain OR 0 stars in comparability domain OR 0 or 1 stars

Appendix 3. Characteristics of included studies.

Study	Study size	Mean age* (range)**	Sex (% male)	Exposure	Outcome	Study type	Location of study	Follow up period (months)	Covariates included
Auf et al. (2018)	39,718	14.5 (12-19)	N/A	Ever-vapers	Ever smoking	Cross-sectional	USA	24	Age, race/ethnicity, gender, peer influence, household e-cigarette use and household use of outcome product
Barrington-Trimis et al. (2018)	6,258	N/A (grade 9-12)	N/A	Ever-vapers	Ever, frequent, and infrequent smoking	Longitudinal	USA	6-18	Gender, race/ethnicity, baseline grade in high school and study (random effect for school)
Best et al. (2018)	2,125	14.4 (11-12)	N/A	Ever-vapers	Ever smoking	Longitudinal	UK	12	Sex, age, ethnicity, and school
Conner et al. (2018)	1,726	13.18 (13-14)	48%	Ever-vapers	Ever smoking	Longitudinal	UK	12	Sex, family smoking, friends' smoking, intentions, attitudes, norms, perceived behavioural control, self-efficacy, and free school meals
East et al. (2018)	1,152	N/A (11-18)	46%	Ever-vapers	Ever smoking	Longitudinal	UK	4-6	Age, gender, school performance, problem behaviour, monthly alcohol use, smoking susceptibility, e-cigarette susceptibility, some friends smoke, some friends use e-cigarettes, at least one parent smokes, at least one parent uses e-cigarettes, sibling(s) smoke, sibling(s) use e-cigarettes and perceived public approval of e-cigarettes.
Hammond et al. (2017)	17,318	N/A (grade 9-12)	47%	Current vapers	Ever smoking	Longitudinal	Canada	12	Age, sex, race/ethnicity, and spending money

Leventhal et al. (2015)	2,530	14.1 (grade 9)	47%	Ever-vapers	Recent smoking	Longitudinal	USA	18	Age, sex, race/ethnicity, parental education, family living situation, family history of smoking, peer smoking, depressive symptoms, impulsivity, use of non–nicotine or tobacco substances, delinquent behaviour, susceptibility to smoking, and smoking outcome expectancies
Loukas et al. (2018)	2,558	19.71 (18-25)	33%	Ever-vapers	Ever smoking	Longitudinal	USA	20	Sex, age, race, type of college attended, susceptibility to smoking, family-of-origin tobacco use, friend cigarette use, and other tobacco use
Lozano et al. (2017)	4,695	N/A (11-13+)	48%	Ever-vapers	Ever smoking	Longitudinal	Mexico	18	Age, sex, parental education, parent smoker, sibling smoker, smoking among close friends, sensation seeking, trial of alcohol, trial of drugs, and internet tobacco product advertising
Miech et al. (2017)	347	Grade 12 (N/A)	44%	Current vapers	Ever smoking	Longitudinal	USA	12	Sex, race, parental education, baseline levels of marijuana use and binge drinking.
Morgenstern et al. (2018)	4,163	15.61 (14-18)	N/A	Ever-vapers	Ever smoking	Longitudinal	Germany	6	Sex, age, federal state, school type, migration background, school leaving qualification of parents, SES, sensation seeking, impulsivity, anxiety sensitivity, hopelessness, extraversion, agreeableness, conscientiousness, neuroticism, openness, alcohol ever, binge drinking ever, cannabis ever, other illegal drugs ever and participation in the “Keep a Clear Head” program.
Primack et al. (2015)	728	N/A (16-26)	46%	Ever-vapers	Ever smoking	Longitudinal	USA	12	Sex, age, race/ethnicity, maternal education level, sensation seeking,

									parental smoking, and smoking among close friends
Primack et al. (2018)	1,506	N/A (18-30)	39%	Ever-vapers	Ever smoking	Longitudinal	USA	12	Age, sex, race, ethnicity, education level, self-esteem, sensation seeking, rebelliousness, yearly household income, living situation and relationship status
Spindle et al. (2017)	2,316	18.5 (N/A)	38%	Ever and current vapers	Current smoking	Longitudinal	USA	12	Sex, age, race/ethnicity, depression, anxiety, negative urgency, positive urgency, lack of premeditation, lack of perseverance, sensation seeking, stressful life events, peer deviance, and other tobacco use
Treur et al. (2018)	6,819	13.8 (11-17)	52%	Ever-vapers	Ever smoking	Longitudinal	Netherlands	6	Age, sex, educational attainment and composite score of smoking propensity
Watkins et al. (2018)	10,348	14.3 (12-17)	51%	Ever and current vapers	Current smoking	Longitudinal	USA	12	Sex, age, race/ethnicity, parental educational, urban residence, sensation seeking, alcohol use, living with tobacco user, frequency of noticing of tobacco warnings, receptivity to tobacco advertising, and season
Wills et al. (2017)	1,141	14.7 (14-16)	47%	Ever-vapers	Ever smoking	Longitudinal	USA	12	Age, sex, race/ethnicity, family structure, parental education, parental support, parental monitoring, sensation seeking, rebelliousness, and clustering within school

*Age reported in years except where grade is stated. Grade reported where it was provided in the study, but actual age was not stated. **At baseline.

Appendix 4. E-cigarette questionnaire items and possible responses.

Variable	Question	Answer
Vaped as a smoking cessation aid	Have you ever used [an e-cigarette] to help you stop smoking?	Tick box response
Ever use	Have you ever used/vaped an electronic cigarette (e-cigarette) or other vaping device?	Yes No
Age of use	How old were you when you first used an electronic cigarette or other vaping device?	Free response
Current use	Do you currently use/vape electronic cigarettes or other vaping devices?	Yes No
Frequency of use	How often did you use/have you used electronic cigarettes/vaping devices?	At least once a day At least once a week At least once a month Less than once a month
Duration of use	How long did you use/have you used electronic cigarettes/vaping devices for?	Less than 1 month 1-3 months 4-6 months 7 months to 1 year 1-2 years More than 2 years
Device type	What type of electronic cigarette/vaping device do you use most often/have you used in the past?	A disposable electronic cigarette or vaping device (non-rechargeable) An electronic cigarette or vaping device that uses replaceable pre-filled cartridges (rechargeable) An electronic cigarette or vaping device with a tank that you refill with liquids (rechargeable) A modular system that you refill with liquids (you use your own combination of separate devices: batteries, atomizers etc.) Rebuildable dripping atomiser (RDA) Other (e.g. e-pipe, e-cigar)
Reasons for use	What are/were your reasons for using electronic cigarettes/vaping devices? Please cross all that apply.	To help me quit smoking To help me cut down on the number of cigarettes I smoke To help me with cravings in situations where I cannot smoke e.g. travel, indoors Pleasure Curiosity Friends use them Other

Smoking just prior to e-cigarette use	Did you smoke tobacco regularly just before you started using electronic cigarettes/vaping devices?	Yes No

Appendix 5. Details of imputed variables.

Variable	Type of variable (number of categories if categorical)	Number (%) with missing data	Regression model used to predict missing data in this variable
Outcome variable			
Use status	Categorical (4)	0 (0%)	N/A
Exposure variables			
To quit	Binary	0 (0%)	N/A
To cut down	Binary	0 (0%)	N/A
To curb cravings	Binary	0 (0%)	N/A
For pleasure	Binary	0 (0%)	N/A
Curiosity	Binary	0 (0%)	N/A
Friends used them	Binary	0 (0%)	N/A
Covariates			
Sex	Binary	0 (0%)	N/A
Parental SEP at birth	Binary	75 (44%)	Logistic regression
Ethnicity	Binary	62 (9%)	Logistic regression
Age at 23+ questionnaire completion	Continuous	7 (1%)	Linear regression
Auxiliary variables (i.e., variables not included in the analysis model)			
Type of device used	Categorical (3)	0 (0%)	N/A
Maternal smoking in pregnancy	Binary	46 (7%)	Logistic regression
BMI	Continuous	121 (18%)	Linear regression
Cannabis use	Binary	207 (31%)	Logistic regression
Other drug use	Binary		Logistic regression
AUDIT score	Continuous	222 (33%)	Linear regression
Ever smoked at age 20	Binary	202 (30%)	Logistic regression
Gambling problems	Binary	228 (34%)	Logistic regression
Condom use	Binary	298 (45%)	Logistic regression
Anxiety	Continuous	256 (38%)	Linear regression
Depressed mood	Binary	262 (39%)	Logistic regression
Employment and education status	Binary	266 (40%)	Logistic regression
Parenthood status	Binary	251 (38%)	Logistic regression
Ever smoked by 17 years	Binary	189 (28%)	Logistic regression
Ever smoked by 18 years	Binary	308 (46%)	Logistic regression
Ever smoked by 20 years	Binary	202 (30%)	Logistic regression
Ever smoked by 21 years	Binary	259 (39%)	Logistic regression
Ever smoked by 22 years	Binary	143 (21%)	Logistic regression
Current smoker at 18 years	Binary	366 (55%)	Logistic regression
Current smoker at 21 years	Binary	276 (41%)	Logistic regression
Current smoker at 22 years	Binary	143 (21%)	Logistic regression
Ever used an e-cigarette by 22 years	Binary	145 (22%)	Logistic regression
Parental SEP at 8 months	Binary	397 (59%)	Logistic regression
Parental SEP at 2 years	Binary	376 (56%)	Logistic regression
Parental SEP at 3 years	Binary	412 (62%)	Logistic regression

Parental SEP at 4 years	Binary	374 (56%)	Logistic regression
Paternal ethnic group	Binary	94 (14%)	Logistic regression
Maternal ethnic group	Binary	156 (23%)	Logistic regression
Paternal ethnic group (reported by mother)	Binary	210 (31%)	Logistic regression
Number of cigarettes mother smoked per day (32 weeks gestation)	Continuous	106 (16%)	Linear regression
Age YP completed 21+ questionnaire	Continuous	248 (37%)	Linear regression
21+ completion date based on date received	Binary	248 (37%)	Logistic regression
Age YP completed 22+ questionnaire	Continuous	140 (21%)	Linear regression
22+ completion date based on date received	Binary	140 (21%)	Logistic regression
BMI at 8	Continuous	381 (57%)	Linear regression
BMI at 16	Continuous	312 (47%)	Linear regression
YP tried cannabis by 15 years	Binary	198 (30%)	Logistic regression
YP tried cannabis by 16 years	Binary	187 (28%)	Logistic regression
YP tried cannabis by 17 years	Binary	272 (41%)	Logistic regression
YP tried cannabis by 18 years	Binary	308 (46%)	Logistic regression
YP tried cannabis by 22 years	Binary	145 (22%)	Logistic regression
Tried other drugs by 16 since 15 - aerosols	Binary	200 (30%)	Logistic regression
Tried other drugs by 16 since 15 - gas	Binary	200 (30%)	Logistic regression
Tried other drugs by 16 since 15 - glue	Binary	198 (30%)	Logistic regression
Tried other drugs by 16 since 15 - solvents	Binary	202 (30%)	Logistic regression
Tried other drugs by 16 since 15 - poppers	Binary	191 (29%)	Logistic regression
Tried other drugs by 16 since 15 - amphetamines	Binary	190 (28%)	Logistic regression
Tried other drugs by 16 since 15 - ecstasy	Binary	188 (28%)	Logistic regression
Tried other drugs by 16 since 15 - LSD	Binary	191 (29%)	Logistic regression
Tried other drugs by 16 since 15 - magic mushrooms	Binary	190 (28%)	Logistic regression
Tried other drugs by 16 since 15 - cocaine	Binary	191 (29%)	Logistic regression
Tried other drugs by 16 since 15 - crack	Binary	192 (29%)	Logistic regression
Tried other drugs by 16 since 15 - heroin	Binary	193 (29%)	Logistic regression
Tried other drugs by 16 since 15 - ketamine	Binary	191 (29%)	Logistic regression
Tried other drugs by 16 since 15 - steroids	Binary	192 (29%)	Logistic regression
Number of illicit drugs used by 22 years	Binary	168 (25%)	Logistic regression
Ever drank alcohol by 16 years	Binary	189 (28%)	Logistic regression
Ever drank alcohol by 17 years		266 (40%)	
AUDIT score at 17 years	Continuous	272 (41%))	Linear regression
Ever drank alcohol by 18 years	Binary	307 (46%)	Logistic regression
Ever drank alcohol by 20 years	Binary	200 (30%)	Logistic regression
Ever drank alcohol by 22 years	Binary	144 (22%)	Logistic regression
AUDIT score at 22 years	Binary	158 (24%)	Logistic regression
Alcohol dependence (DSM-4)	Binary	159 (24%)	Logistic regression
Alcohol abuse (DSM-4)	Binary	152 (23%)	Logistic regression
Gambled by 18 years	Binary	290 (43%)	Logistic regression
Had sexual intercourse in last year at 15 years	Binary	492 (74%)	Logistic regression
Used condom when last had sex at 15 years	Binary	564 (84%)	Logistic regression
Had sexual intercourse in last year at 23 years	Binary	7 (1.05%)	Logistic regression

Used condom when last had sex at 23 years	Categorical (3)	18 (3%)	Multinomial logistic regression
Total CIS-r score at 18 years	Continuous	231 (35%)	Linear regression
Anxiety score at 17 years	Continuous	231 (35%)	Linear regression
YP been sad/miserable/tearful	Binary	192 (29%)	Logistic regression
YP been sad/miserable/tearful for 3+ hours regularly	Binary	504 (75%)	Logistic regression
Depression score at 17 years	Continuous	231 (35%)	Linear regression
YP in full time education at 16 years	Binary	190 (28%)	Logistic regression
YP working part time at 16 years	Binary	186 (28%)	Logistic regression
YP working full time at 16 years	Binary	186 (28%)	Logistic regression
YP in full time education or employment at 20 years	Binary	208 (31%)	Logistic regression
YP become a parent by 16 years	Binary	196 (29%)	Logistic regression
YP become a parent by 20 years	Binary	195 (29%)	Logistic regression

YP = young person; BMI = body mass index; AUDIT = alcohol use disorder identification test; CIS-r = computerised interview scale - revised

Appendix 6. Associations between reasons for vaping by 23 years and current vaping at 24 years among ever vapers and ever smokers.

Reason for vaping by 23 years	Unadjusted (n=668)			Adjusted (n=578)		
	OR	95% CI	*p*-value	OR	95% CI	*p*-value
To quit smoking	3.43	2.24, 5.23	<.001	3.97	2.50, 6.30	<.001
To cut down	2.80	1.82, 4.31	<.001	3.10	1.95, 4.93	<.001
To curb cravings	4.34	2.58, 7.29	<.001	4.95	2.82, 8.68	<.001
Pleasure	3.20	2.02, 5.07	<.001	3.08	1.86, 5.10	<.001
Curiosity	0.42	0.27, 0.65	<.001	0.39	0.24, 0.62	<.001
Friends used them	0.62	0.36, 1.07	.084	0.51	0.28, 0.94	.031

Note: The analyses were restricted to individuals who ever smoked and ever vaped. Adjusted analyses adjusted for demographic factors (sex, ethnicity, socioeconomic position, and age in months at 23-year questionnaire). OR = Odds ratio.

Appendices

Appendix 7. Associations between reasons for vaping by 23 years and current smoking at 24 years among ever vapers and regular smokers prior to vaping.

Reason for vaping by 23 years	Unadjusted (n=412)			Adjusted (n=360)		
	OR	95% CI	*p*-value	OR	95% CI	*p*-value
To quit smoking	0.49	0.32, 0.77	.002	0.48	0.30, 0.77	.003
To cut down	1.68	1.06, 2.65	.027	1.61	0.98, 2.64	.058
To curb cravings	0.91	0.52, 1.60	.744	0.85	0.46, 1.58	.614
Pleasure	0.92	0.50, 1.70	.797	0.75	0.39, 1.43	.378
Curiosity	1.68	1.06, 2.65	.027	1.72	1.04, 2.84	.034
Friends used them	1.76	0.94, 3.30	.077	1.74	0.91, 3.33	.092

Note: The analyses were restricted to individuals who ever smoked and had smoked regularly just before they started vaping. Adjusted analyses adjusted for demographic factors (sex, ethnicity, socioeconomic position, and age in months at 23-year questionnaire). OR = Odds ratio.

Appendices

Appendix 8. List of SNPs included in analyses

SNP rsID	Chromosome	Position	Associated phenotype	Proxy SNP rsID	Included in Cotinine MR	Included in CPD MR	Included in MVMR main analysis	Included in MVMR supplementary analysis
rs10851907	15	78915864	Cotinine		Yes	No	Yes	Yes
rs77107237	4	69746647	Cotinine		Yes	No	Yes	Yes
rs57064725	15	78540694	Cotinine		No	No	No	No
rs4887074	15	78952110	Cotinine		Yes	No	Yes	No
rs74386627	15	79099986	Cotinine		Yes	No	Yes	No
rs7971121	12	129977245	Cotinine		Yes	No	Yes	No
rs2814672	6	63134804	Cotinine		Yes	No	Yes	No
rs76474922	15	78884553	Cotinine		Yes	No	Yes	No
rs11745112	5	685849	Cotinine		Yes	No	Yes	No
rs1813379	16	84437335	Cotinine		Yes	No	Yes	No
rs10744625	12	3868168	Cotinine		Yes	No	Yes	No
rs11264100	1	35591626	Cigarettes per day	rs2971426	No	Yes	Yes	Yes
rs2072659	1	154548521	Cigarettes per day		No	Yes	Yes	Yes
rs34973462	1	175993820	Cigarettes per day		No	Yes	Yes	Yes
rs7599488	2	60718347	Cigarettes per day		No	Yes	Yes	Yes
rs78408772	2	62710608	Cigarettes per day		No	Yes	Yes	Yes
rs10204824	2	148372720	Cigarettes per day		No	Yes	Yes	Yes
rs2084533	3	16872929	Cigarettes per day		No	Yes	Yes	Yes

rs7431710	3	48935583	Cigarettes per day	rs11705737	No	Yes	Yes	Yes
rs2236951	3	50421081	Cigarettes per day		No	Yes	Yes	Yes
rs699165	3	136224697	Cigarettes per day		No	Yes	Yes	Yes
rs28813180	3	158083918	Cigarettes per day		No	Yes	Yes	Yes
rs1024323	4	3006043	Cigarettes per day		No	Yes	Yes	Yes
rs11940255	4	67086288	Cigarettes per day		No	Yes	Yes	Yes
rs10454798	4	67980830	Cigarettes per day		No	Yes	Yes	Yes
rs7766641	6	26184102	Cigarettes per day		No	Yes	Yes	Yes
rs215600	7	32333642	Cigarettes per day		No	Yes	Yes	Yes
rs62447179	7	50339609	Cigarettes per day		No	Yes	Yes	Yes
rs2741351	8	27418040	Cigarettes per day		No	Yes	Yes	Yes
rs73229090	8	27442127	Cigarettes per day		No	Yes	Yes	Yes
rs13253502	8	42442018	Cigarettes per day		No	Yes	Yes	Yes
rs4236926	8	42578059	Cigarettes per day		No	Yes	Yes	Yes
rs790564	8	64604218	Cigarettes per day		No	Yes	Yes	Yes

rs75596189	9	136468701	Cigarettes per day		No	Yes	Yes	Yes
rs3025383	9	136502369	Cigarettes per day		No	Yes	Yes	Yes
rs7951365	11	16377044	Cigarettes per day		No	Yes	Yes	Yes
rs10742683	11	43667625	Cigarettes per day		No	Yes	Yes	Yes
rs113001570	11	46737412	Cigarettes per day	rs77920005	No	Yes	Yes	Yes
rs7125588	11	113436072	Cigarettes per day		No	Yes	Yes	Yes
rs11846838	14	104184737	Cigarettes per day		No	Yes	Yes	Yes
rs1115019	15	57141231	Cigarettes per day		No	Yes	Yes	Yes
rs632811	15	59155050	Cigarettes per day		No	Yes	Yes	Yes
rs182317	15	89943601	Cigarettes per day		No	Yes	Yes	Yes
rs1592485	16	52093549	Cigarettes per day		No	Yes	Yes	Yes
rs12924872	16	69552215	Cigarettes per day		No	Yes	Yes	Yes
rs258321	16	89756473	Cigarettes per day		No	Yes	Yes	Yes
rs4144686	18	53251725	Cigarettes per day		No	Yes	Yes	Yes
rs4485470	18	62125063	Cigarettes per day		No	Yes	Yes	Yes

rs59208569	19	4044424	Cigarettes per day		No	Yes	Yes	Yes
rs143200968	19	41338847	Cigarettes per day		No	Yes	Yes	Yes
rs56113850	19	41353107	Cigarettes per day		No	Yes	Yes	Yes
rs6078373	20	11863500	Cigarettes per day		No	Yes	Yes	Yes
rs1737894	20	31054702	Cigarettes per day		No	Yes	Yes	Yes
rs2273500	20	61986949	Cigarettes per day		No	Yes	Yes	Yes
rs7281463	21	40520783	Cigarettes per day		No	Yes	Yes	Yes
rs8192726	19	41354496	Cigarettes per day	rs76112798	No	Yes	No	No
rs8040868	15	78911181	Cigarettes per day		No	Yes	No	No
rs3743063	15	79065171	Cigarettes per day		No	Yes	No	No
rs28681284	15	78908565	Cigarettes per day		No	Yes	No	No
rs10519203	15	78814046	Cigarettes per day		No	Yes	No	No
rs12438181	15	78812098	Cigarettes per day		No	Yes	No	No
rs28438420	15	78836288	Cigarettes per day		No	Yes	No	No
rs146009840	15	78906177	Cigarettes per day		No	Yes	No	No

rs72740955	15	78849779	Cigarettes per day	No	Yes	No	No
rs4886550	15	78243579	Cigarettes per day	No	Yes	No	No

Appendix 9. Tests of the weighted regression dilution (I^2_{GX}), instrument validity and heterogeneity among ever smokers in MR analyses.

Exposure	Outcome	Q	I^2_{GX} (weighted)	mF
Cotinine ($5x10^{-6}$)	BMI	22.00	0.743	30.67
	COPD	22.03	0.793	
	FEV-1	27.71	0.747	
	FVC	10.27	0.747	
	CHD	23.60	0.742	
	Heart Rate	16.47	0.743	
Cotinine ($5x10^{-8}$)	BMI	8.02	0.573	61.29
	COPD	13.93	0.629	
	FEV-1	3.83	0.577	
	FVC	0.23	0.577	
	CHD	2.10	0.585	
	Heart Rate	5.58	0.573	
CPD	BMI	270.37	0.965	68.44
	COPD	144.80	0.965	
	FEV-1	258.10	0.965	
	FVC	184.00	0.965	
	CHD	138.82	0.965	
	Heart Rate	165.92	0.965	

Note: Cotinine (5×10^{-6}) refers to MR analysis in which cotinine SNPs were selected for inclusion based on a threshold of $p < 5 \times 10^{-6}$. Cotinine ($5x10^{-8}$) refers to MR analysis in which cotinine SNPs were selected for inclusion based on a threshold of $p < 5\times10^{-8}$. CPD = cigarettes per day. BMI = body mass index; COPD = chronic obstructive pulmonary disease; FEV-1 = forced expiratory volume in 1 second; FVC = forced vital capacity; CHD = chronic heart disease; mF = mean F-statistic. mF tests for instrument validity. The Cochran *Q* statistic tests for instrument strength and validity in the two sample summary data setting.

Appendix 10. Tests of the weighted regression dilution (I^2_{GX}), instrument validity and heterogeneity among current smokers in MR analyses.

Exposure	Outcome	Q	I^2_{GX} (weighted)	mF
Cotinine ($5x10^{-6}$)	BMI	41.22	0.760	30.67
	COPD	19.91	0.847	
	FEV-1	12.39	0.766	
	FVC	2.29	0.766	
	CHD	13.16	0.733	
	Heart Rate	16.60	0.757	
Cotinine ($5x10^{-8}$)	BMI	19.10	0.531	61.29
	COPD	14.09	0.702	
	FEV-1	4.58	0.517	
	FVC	0.24	0.516	
	CHD	3.27	0.557	
	Heart Rate	4.42	0.530	
CPD	BMI	217.21	0.965	68.44
	COPD	99.42	0.965	
	FEV-1	177.36	0.965	
	FVC	114.84	0.965	
	CHD	76.32	0.965	
	Heart Rate	135.87	0.965	

Note: Cotinine ($5x10^{-6}$) refers to MR analysis in which cotinine SNPs were selected for inclusion based on a threshold of $p < 5\times10^{-6}$. Cotinine ($5x10^{-8}$) refers to MR analysis in which cotinine SNPs were selected for inclusion based on a threshold of $p < 5\times10^{-8}$. CPD = cigarettes per day. BMI = body mass index; COPD = chronic obstructive pulmonary disease; FEV-1 = forced expiratory volume in 1 second; FVC = forced vital capacity; CHD = chronic heart disease. mF = mean F-statistic. The Cochran Q statistic tests for instrument strength and validity in the two sample summary data setting. mF tests for instrument validity.

Appendix 11. Univariable Mendelian randomisation analysis of cotinine and smoking-related health outcomes among ever and current smokers (n = 3 SNPs).

Health outcome	MR method	Ever smokers			Current smokers		
		Effect	95% CI	*p*-value	Effect	95% CI	*p*-value
BMI	IVW	-0.08	-0.31, 0.14	0.469	-0.78	-1.52, -0.04	0.039
	MR-Egger (SIMEX)	0.25	-11.79, 12.29	0.967	-1.58	-43.76, 40.59	0.941
	W. median	-0.13	-0.26, 0.00	0.045	-1.04	-1.36, -0.71	<0.001
	W. mode	-0.16	-0.32, 0.00	0.057	-1.05	-1.40, -0.70	<0.001
COPD	IVW	0.54	-0.24, 1.33	0.175	0.64	-0.83, 2.10	0.393
	MR-Egger (SIMEX)	-2.76	-34.00, 28.48	0.863	-4.43	-47.27, 38.42	0.840
	W. median	0.56	0.18, 0.94	0.004	0.75	0.07, 1.43	0.030
	W. mode	0.46	-0.12, 1.05	0.121	0.80	-0.22, 1.83	0.124
FEV-1	IVW	-0.07	-0.10, -0.04	<0.001	-0.14	-0.22, -0.07	<0.001
	MR-Egger (SIMEX)	-0.03	-1.48, 1.42	0.970	-0.14	-2.85, 2.57	0.918
	W. median	-0.08	-0.10, -0.05	<0.001	-0.17	-0.23, -0.11	<0.001
	W. mode	-0.08	-0.11, -0.05	<0.001	-0.17	-0.25, -0.09	<0.001
FVC	IVW	-0.05	-0.07, -0.02	<0.001	-0.10	-0.16, -0.04	<0.001
	MR-Egger (SIMEX)	-0.02	-0.40, 0.36	0.918	-0.15	-1.37, 1.07	0.810
	W. median	-0.05	-0.08, -0.02	0.001	-0.11	-0.18, -0.04	0.001
	W. mode	-0.05	-0.08, -0.01	0.008	-0.11	-0.19, -0.03	0.008
CHD	IVW	0.02	-0.06, 0.10	0.628	0.07	-0.14, 0.27	0.519
	MR-Egger (SIMEX)	0.22	-2.86, 3.30	0.888	0.58	-7.84, 8.99	0.893
	W. median	0.04	-0.05, 0.12	0.379	0.12	-0.06, 0.30	0.203
	W. mode	0.05	-0.06, 0.15	0.393	0.14	-0.09, 0.37	0.226
Heart rate	IVW	0.52	0.05, 0.98	0.030	1.52	0.62, 2.43	<0.001
	MR-Egger (SIMEX)	-1.20	-7.87, 5.46	0.723	-1.37	-14.19, 11.45	0.834
	W. median	0.54	0.19, 0.89	0.003	1.60	0.81, 2.39	<0.001
	W. mode	0.53	0.06, 1.00	0.026	1.64	0.62, 2.66	0.002

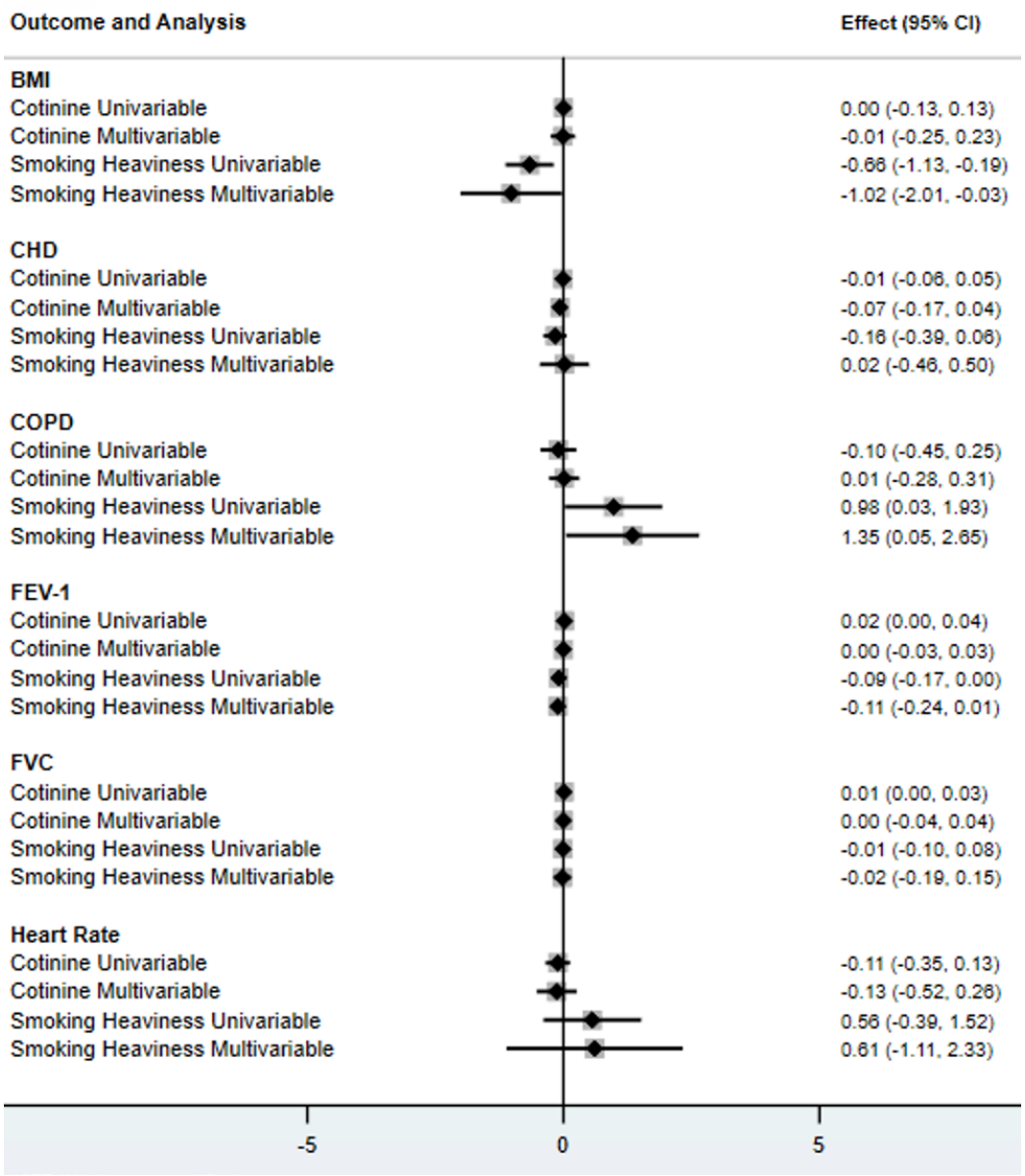
Note: Single nucleotide polymorphisms (SNPs) identified as conditionally independently associated with cotinine in a GWAS were selected for inclusion. BMI = body mass index; COPD = chronic obstructive pulmonary disease; FEV-1 = forced expiratory volume in 1 second; FVC = forced vital capacity; CHD = chronic heart disease; IVW = inverse variance weighted; W. median = weighted median; W. mode = weighted mode. Effects are betas for continuous variables (BMI, FEV-1, FVC and heart rate) and log odds ratios for binary outcomes (COPD and CHD).

Appendix 12. MR-Egger test of directional pleiotropy among ever smokers for cotinine ($p < 5x10^{-8}$).

	Ever smokers			Current smokers		
Outcome	Intercept	95% CI	*p*-value	Intercept	95% CI	*p*-value
BMI	-0.06	-2.85, 2.73	0.967	0.21	-9.34, 9.75	0.966
COPD	0.67	-6.34, 7.67	0.852	1.00	-8.53, 10.53	0.837
FEV-1	-0.01	-0.34, 0.32	0.961	0.00	-0.64, 0.65	0.993
FVC	-0.01	-0.10, 0.08	0.899	0.01	-0.28, 0.30	0.942
CHD	-0.05	-0.74, 0.65	0.895	-0.12	-2.04, 1.80	0.904
Heart Rate	0.35	-0.91, 1.60	0.588	0.58	-1.83, 2.99	0.638

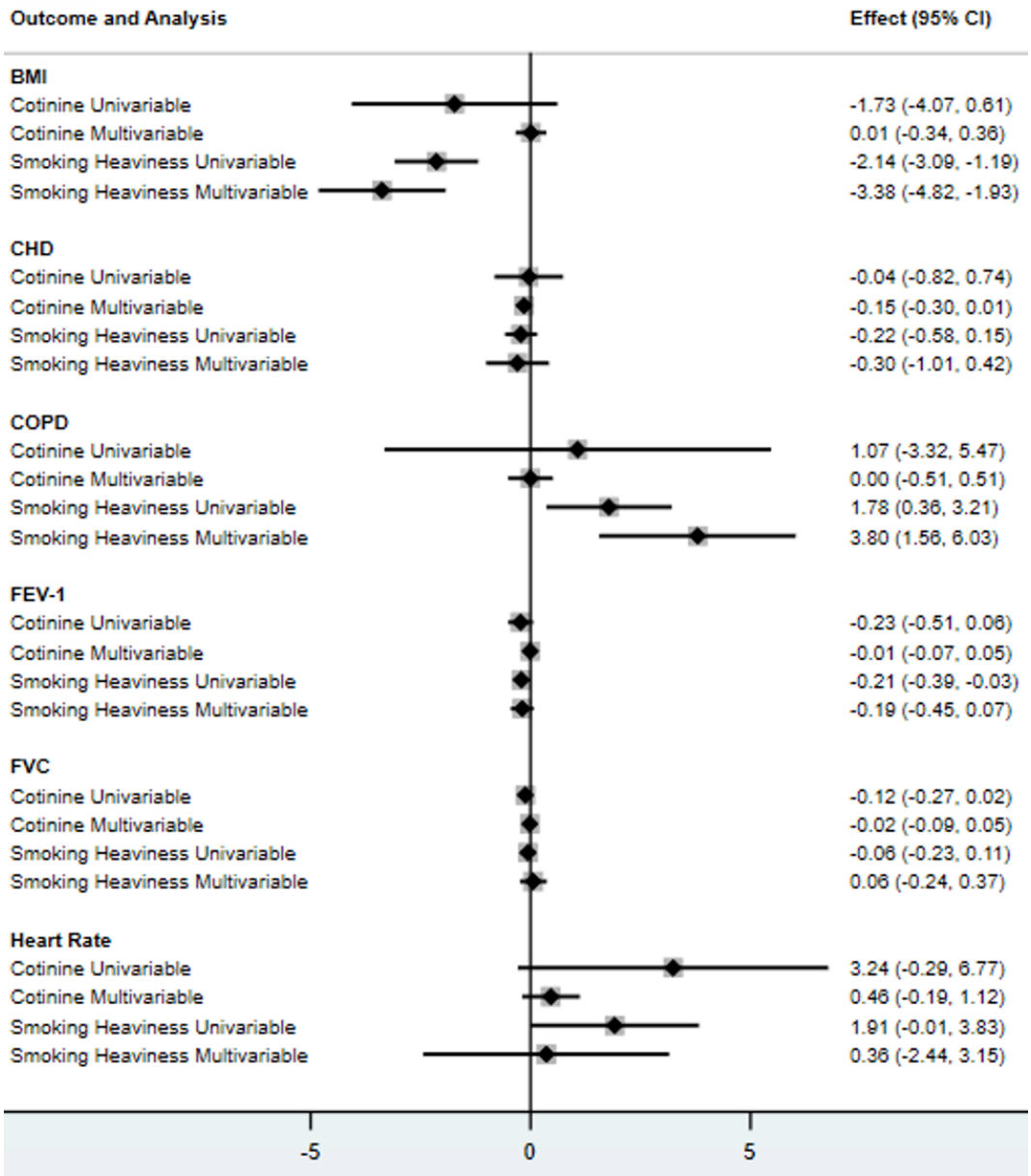
Note: Cotinine ($5x10^{-8}$) refers to MR analysis in which cotinine SNPs were selected for inclusion based on a threshold of $p < 5\times10^{-8}$. BMI = body mass index; COPD = chronic obstructive pulmonary disease; FEV-1 = forced expiratory volume in 1 second; FVC = forced vital capacity; CHD = chronic heart disease.

Appendix 13. Univariable and multivariable Mendelian randomisation Egger analysis of cotinine and smoking heaviness (cigarettes per day) and smoking-related health outcomes among ever smokers (n = 54 SNPs).



Appendices

Appendix 14. Univariable and multivariable Mendelian randomisation Egger analysis of cotinine and smoking heaviness (cigarettes per day) and smoking-related health outcomes among current smokers (n = 54 SNPs).



Appendix 15. Multivariable Mendelian randomisation analysis of cotinine and cigarettes per day (CPD) and smoking-related health outcomes among ever, current, former and never smokers (n = 46 SNPs).

Outcome	Smoking Status	Cotinine			CPD			*Q*
		Effect	95% CI	*p*-value	Effect	95% CI	*p*-value	
BMI	Ever	-0.02	-0.34, 0.30	0.897	0.10	-0.56, 0.75	0.770	251.33
	Current	-0.01	-0.45, 0.43	0.964	-1.61	-2.52, -0.71	<0.001	103.67
	Former	0.01	-0.31, 0.33	0.952	0.59	-0.07, 1.26	0.077	201.42
	Never	-0.07	-0.35, 0.21	0.612	0.97	0.39, 1.54	0.001	225.98
COPD	Ever	0.17	-0.20, 0.54	0.357	2.23	1.48, 2.98	<0.001	44.81
	Current	-0.22	-0.80, 0.35	0.439	3.95	2.79, 5.12	<0.001	33.67
	Former	0.29	-0.18, 0.77	0.220	1.46	0.49, 2.44	0.004	52.94
	Never	0.36	-0.30, 1.03	0.277	-1.69	-3.05, -0.33	0.016	39.79
FEV-1	Ever	-0.02	-0.06, 0.02	0.429	-0.22	-0.30, -0.14	<0.001	100.76
	Current	-0.05	-0.13, 0.03	0.201	-0.35	-0.51, -0.19	<0.001	75.89
	Former	0.00	-0.04, 0.04	0.848	-0.19	-0.27, -0.10	<0.001	92.31
	Never	0.01	-0.03, 0.04	0.733	-0.04	-0.11, 0.03	0.275	92.47
FVC	Ever	0.00	-0.06, 0.05	0.907	-0.21	-0.32, -0.10	<0.001	121.64
	Current	-0.05	-0.15, 0.05	0.302	-0.25	-0.45, -0.05	0.014	78.73
	Former	0.01	-0.05, 0.06	0.779	-0.20	-0.31, -0.08	0.001	110.34
	Never	0.01	-0.04, 0.05	0.802	-0.05	-0.15, 0.04	0.270	114.94
CHD	Ever	-0.05	-0.19, 0.08	0.431	0.36	0.08, 0.63	0.013	100.41
	Current	0.01	-0.20, 0.22	0.947	0.29	-0.14, 0.72	0.177	53.53
	Former	-0.07	-0.22, 0.08	0.360	0.37	0.06, 0.69	0.021	99.73
	Never	-0.13	-0.24, -0.02	0.023	0.09	-0.14, 0.31	0.437	44.97
Heart rate	Ever	0.28	-0.25, 0.81	0.289	1.71	0.64, 2.79	0.002	104.04
	Current	1.04	0.22, 1.85	0.014	3.08	1.41, 4.74	<0.001	51.07
	Former	-0.04	-0.61, 0.53	0.890	1.31	0.14, 2.48	0.029	103.47
	Never	0.03	-0.53, 0.59	0.917	-0.09	-1.24, 1.06	0.875	158.66

Note: A *p*-threshold of 5×10^{-8} was used to determine the single nucleotide polymorphisms (SNPs) associated with cigarettes per day (CPD) and cotinine. BMI = body mass index; COPD = chronic obstructive pulmonary disease; FEV-1 = forced expiratory volume in 1 second; FVC = forced vital capacity; CHD = chronic heart disease. Effects are betas for continuous variables (BMI, FEV-1, FVC and HR) and log odds ratios for binary outcomes (COPD and CHD). The Cochran *Q* statistic tests for instrument strength and validity in the two sample summary data setting.

Appendix 16. Contraindications for nicotine or placebo use.

Contraindication
Panic attacks
Diagnosed and medicated asthma attacks
Stomach ulcers
Liver or kidney disease
Overactive thyroid gland
Diabetes
Prone to allergic reactions*
Prone to seizures
Taking medication (excluding the contraceptive pill)

Note: *Participants were provided with a full list of ingredients for the nicotine and placebo sprays.