Chronic diseases are increasingly important in the developed world. It has been proposed that adverse exposures during pregnancy can harm the fetus and subsequently increase the susceptibility to develop chronic diseases later in life. This thesis aimed at using the Developmental Origin of Health and Disease paradigm as a framework to conduct environmental health risk assessment to estimate disease burden associated with maternal smoking as an example of a developmental exposure. The effect of maternal smoking during early and late pregnancy on body size and proportions at birth were investigated. Potential loss of healthy life in the Finnish population attributable to prenatal exposure to tobacco smoke was quantified.
MATERNAL SMOKING, BIRTH OUTCOMES AND LATER LIFE HEALTH

ESTIMATION OF THE DEVELOPMENTAL ORIGIN OF DISEASE BURDEN
Isabell Katharina Rumrich

MATERNAL SMOKING, BIRTH OUTCOMES AND LATER LIFE HEALTH

ESTIMATION OF THE DEVELOPMENTAL ORIGIN OF DISEASE BURDEN

Publications of the University of Eastern Finland
Dissertations in Forestry and Natural Sciences
No 365

University of Eastern Finland
Kuopio
2020

Academic dissertation
To be presented by permission of the Faculty of Science and Forestry for public examination in the Auditorium CA102 in the Canthia Building at the University of Eastern Finland, Kuopio, on February, 7, 2020, at 12 o’clock noon.
Author’s address: Isabell Katharina Rumrich
University of Eastern Finland
Depart. of Environmental and Biological Sciences
70210 KUOPIO, FINLAND
And
Finnish Institute for Health and Welfare
Department of Public Health Solutions
70701 KUOPIO, FINLAND
email: isabell.rumrich@thl.fi

Supervisors: Docent Otto Hänninen, Ph.D.
Finnish Institute for Health and Welfare
Department of Public Health Solutions
70701 KUOPIO, FINLAND
email: otto.hanninen@thl.fi

Professor (emer.) Matti Viluksela, Ph.D.
University of Eastern Finland
Depart. of Environmental and Biological Sciences
And
School of Pharmacy/Toxicology
70210 KUOPIO, FINLAND
email: matti.viluksela@uef.fi

Professor (emer.) Kirsi Vähäkangas, Ph.D., MD
University of Eastern Finland
School of Pharmacy/Toxicology
70210 KUOPIO, FINLAND
email: kirsi.vahakangas@uef.fi

Reviewers: Professor Bertil Forsberg, Ph.D.
Umeå University
Department of Public Health and Clinical Medicine
90185 UMEÅ, SWEDEN
email: bertil.forsberg@umu.se

Associate Professor Hely Katariina Laine, Ph.D., MD
Department of Obstetrics
Oslo University Hospital
0424 OSLO, NORWAY
email: h.k.laine@medisin.uio.no

Opponent: Professor Jouni Jaakkola, Ph.D., MD
Center for Environmental and Respiratory Health Research (CERH)
University of Oulu
90014 University of OULU, FINLAND
email: jouni.jaakkola@oulu.fi
ABSTRACT

The relative disease burden of non-communicable diseases is increasing but established risk factors and exposures explain only a fraction of the burden. The paradigm of developmental origin of health and disease (DOHaD) suggests that stress and insults during prenatal development may harm the fetus and increase the susceptibility to diseases later in life. Adverse health effects of tobacco smoke in adults have long been established, but long-term effects of exposure to prenatal maternal smoking have not yet been resolved. Maternal smoking is a readily modifiable risk factor.

The aim of this doctoral dissertation was to use the DOHaD paradigm as a framework to estimate disease burden associated with developmental exposures. Specifically, the objectives were to characterise smoking behaviour during pregnancy; to analyse the effect of maternal smoking during early and late pregnancy on body size and proportions at birth; and to estimate the disease burden in later life of the child attributable to maternal smoking. Additionally, the potential of Finnish Register data for long-term follow-up of a birth cohort was evaluated.

The MATEX birth cohort was established from the Finnish Medical Birth Register including all births from 1987 to 2016 (ca. 1.7 million mother-child pairs). Smoking information was available from 1991 to 2016. The cohort contains information on the mother’s background including maternal smoking and socio-demographics, as well as complications and diagnoses during pregnancy and births. Anthropometric measurements, gestational age, and diagnoses up to 7 days of age are available for the child.

Self-reported smoking during pregnancy fluctuated between 13.8% and 16.3% (mean 15%) between 1991 and 2016 in Finland. However, the fraction of women quitting smoking during the 1st trimester rose from 2% to 7% in the same period. Smoking rates were higher in younger pregnant women and in those with lower socioeconomic status. About 40% of women who smoked during their 1st pregnancy were non-smokers during their 2nd pregnancy.
Smoking during pregnancy was associated with changes in body proportions at birth indicated by a high ponderal index (OR 1.26; 95%CI 1.23-1.28), a low brain-to-body ratio (1.11; 1.07-1.15) and a high head-to-length ratio (1.22; 1.19-1.26). The associations show a stronger reduction in length and in brain size than in weight. The effects were slightly more pronounced for smoking throughout pregnancy than for smoking only during early pregnancy. Although quitting smoking during early pregnancy reduced the risk for preterm birth to background level, the association with generalized non-proportional growth restriction stressed the importance of the period of early prenatal development, especially in brain development, and the limited potential to repair damages later during pregnancy.

In a DOHaD-based chained risk model maternal smoking was linked with intermediate risk factors (low birth weight, preterm birth and childhood overweight) and subsequently with later life disease in the child. Maternal smoking and intermediate risk factors were found to be associated, amongst others, with cardiovascular disease, diabetes, cancer, asthma, and mental health. Accounting for direct and chained associations, roughly 1,200 disability-adjusted life years (DALY) were attributable to maternal smoking in Finland in 2017.

The long tradition of health registers in Finland provides excellent opportunity to study prenatal exposures. Register linkage of the Medical Birth Register with other health registers allows to follow individuals from in utero throughout the whole life course up until death.

In summary, maternal smoking is associated with increased risk for life-long health consequences, contributing to the rise in chronic disease burden. The sensitivity of early prenatal development stresses the importance for smoking cessation before conception to avoid persistent health effects. This work confirms the sensitivity of developmental periods to harmful exposures, here using the example of maternal smoking, which should be considered in comprehensive risk assessment. Furthermore, this work demonstrates that developmental exposures can be used to explain part of previously unknown causes of burden of disease.

National Library of Medicine Classification: QZ 185, WM 290, WQ 200, WQ 210

Medical Subject Headings: Tobacco Smoking; Pregnancy; Pregnant Women; Maternal Exposure; Prenatal Exposure Delayed Effects; Risk Factors; Fetus; Embryonic and Fetal Development; Body Constitution; Body Size; Birth Weight; Premature Birth; Pediatric Obesity; Health; Chronic Disease; Maternal Age; Social Class; Registries; Finland
Rumrich, Isabell Katharina
Raskausajan tupakointi, syntymään liittyvät terveysvasteet ja myöhemi elämä - Arviointi raskausajan vaikutuksista lapsen myöhempään tautitaakkaan

TIIVISTELMÄ

Tarttumattomien tautien tautitaakka on kasvussa kehittyneissä maissa. Vain osa tästä tautitaakasta selittyy tunnetuilla riskitekijöillä. Sikiökautisen ohjelmoitumisen paradigma viittaa siihen, että stressi ja erilaiset vauriot prenataalisen kehityksen aikana voivat vahingoittaa sikiötä ja lisätä herkkyyttä sairauksille myöhemmässä elämässä. Tupakansavun haitalliset vaikutukset on tunnettu jo pitkään, mutta raskausajan tupakoinnin pitkäaikaisvaikutuksista syntyvälle lapselle on toistaiseksi vähän tietoa. Raskausajan tupakointi on riskitekijä, johon on mahdollista vaikuttaa.

Tämän väitöskirjan tavoitteena on käyttää sikiökautisen ohjelmoitumisen paradigmat menkyn kaikista kehitysaikaisista altistumisen myöhemmän tautitaakan arviointiin. Erityisesti tavoitteena on kuvata raskaudenaikaita tupakointikäyttäytymistä; arvioidaan alku- ja loppuraskaudessa tapahtuvan tupakoinnin vaikutuksia syntyvän lapsen syntymäkokoon ja mittasuhteisiin; sekä määritätä raskaudenaikaisen tupakoinnin aiheuttamaa tautitaakkaa lapsen myöhemmässä elämässä. Lisäksi arvioitiin suomalaisten rekisteritietojen käyttöä syntymäkohortin pitkäaikaisessa seurannassa.


Äitien raportoima raskausajan tupakointi vaihteli 13.8 % ja 16.3 % välillä (keskiarvo 15 %) Suomessa vuosina 1991-2016. Tällä aikavälillä naisten tupakoinnin lopettaminen ensimmäisen kolmanneksen aikana lisääntyi 2 %:sta 7 %:iin. Nuoret ja alemmassa sosioekonomisessa asemassa olevat äidit tupakoivat muihin enemmän. Ensiraskauden aikana tupakoivista äideistä noin 40 % eivät tupakoineet toisen raskauden aikana.

Raskausajan tupakointi oli yhteydessä muutoksiin vauvan kehon mittasuhteissa, mikä näkyi kohonneena paino-pituus suhteena (ponderaal-indeksi; OR 1.26, 95%LV 1.23-1.28), alentuneena aivojen tilavuuden ja kehon painon suhteena (OR 1.11; 1.07-1.15) ja kohonneena päänympärys-pituus-suhteena (OR 1.22; 1.19-1.26). Tulos viittaa voimakkaampaan pituuden ja aivojen koon pienememiseen verrattuna painon alenemiseen. Koko raskauden kestäneen tupakoinnin riskisuhde olit selkeämpi kuin pelkän alkuraskauden. Vaikka tupakoinnin lopettaminen alkuraskauden aikana vähensi ennenaiaksen syntymän riskin taustatasolle, yhteys kehon mittasuhteiden muutokseen säilyi viitaten pienempänä aivojen kokoon ja tupakoinnin lopettamisen rajalliseen mahdollisuuteen estää varhaisessa raskaudessa syntyneitä vahinkoja.

Suomessa terveysrekisterit tarjoavat erinomaiset mahdollisuudet sikiöaikaisen altistumisen myöhempien vaikutusten tutkimiseen linkittämällä syntymärekisteri muihin terveysrekistereihin.


Luokitus: QZ 185, WM 290, WQ 200, WQ 210

Yleinen suomalainen ontologia: tupakka; tupakointi; raskaus; äidit; altistuminen; riskitekijät; sikiö; sikiönkehitys; vastasyntyneet; mittasuhteet; koko; syntymäpaimo; ennenaikainen synnytys; ylipaimo; lapsuus; pitkääikaisvaikutukset; terveys; krooniset taudit; rekisterit; Suomi
“I believe that the most important single thing, beyond discipline and creativity is daring to dare.”
— Maya Angelou
ACKNOWLEDGEMENTS

This work was carried out in the Department of Public Health Solutions at the Finnish Institute for Health and Welfare (THL), and the Department of Biological and Environmental Sciences at the University of Eastern Finland (UEF), Kuopio, between 2014 and 2019. This work was conducted under the Environmental Physics, Health and Biology Doctoral programme of UEF.

I owe my warmest gratitude to my outstanding supervisors Docent Otto Hänninen, Prof (emer.) Matti Viluksela and Prof (emer.) Kirsi Vähäkangas for the guidance throughout this project. I am grateful for the opportunity to explore the world of scientific research under their guidance and support. It has taught me so much more than just science. Without their continuous encouragement and believe in me this work would have been impossible.

I wish to express my deepest gratitude to my principal supervisor, Docent Otto Hänninen, for his scientific guidance and support throughout this work. He has guided me through this process, led me in an academic thinking and accuracy as well as supported my research efforts in an international context. I would like to thank him for inspiring philosophical discussions. I appreciate his braveness to step out of well-known territory and leap into register-based epidemiology for this project. I owe my gratitude to my second supervisor, Prof (emer.) Matti Viluksela. His knowledge in environmental health was invaluable for this work. His advice and support were invaluable. I also appreciate his help in all administrative matters. Last, but not least, I owe my gratitude to my third supervisor, Professor (emer.) Kirsi Vähäkangas. Her interest and knowledge in prenatal exposures and tobacco have been inspiring. Without her positive attitude and encouragement this project would have been a lot harder.

My appreciation goes to my pre-examiners of my dissertation manuscript. Thank you, Professor Bertil Forsberg from Umeå University and Associate Professor Hely Katariina Laine from the Oslo University Hospital. Your constructive comments and valuable suggestions were invaluable and have helped and enriched my work.

I would like to extend my sincere appreciation to my co-authors Professor Mika Gissler, MD Hanna de Ruyter, and Docent Heljä-Marja Surcel. Their collaboration and efforts made this work possible.

My owe my gratitude to the Department of Public Health Solutions at THL and the ROKOSTAT-team, especially Esa Ruokokoski and Jonas Sundman, who provided the Linux Cluster environment and implemented the Medical Birth Register linkage. Essential support from the Register Holder by Jouni Meriläinen was provided.

I am fortunate to have the support and encouragement from fellow colleagues from the Finnish Institute of Health and Welfare and the University of Eastern Finland during these past years of my doctoral research work. It is not possible to list all deserving, but amongst them I want to acknowledge Heli Leottomäki Zrim, Antti Korhonen, Jouni Tuomisto, Päivi Meriläinen, Jukka Jokinen, Sanna Tossavainen and
Jouko Tuomisto. For the opportunity to gain experience in teaching, I want to thank Jaana Rysä and my fellow young university teacher colleagues.

I believe I would not be in this situation without the support from my family and friends. It is not possible to list all deserving, but you know who you are. Thank you for making Finland home and thank you for sticking around even with thousands of kilometres between us. You have a special place in this adventure. To my family, thank you for pushing me through the uphills of life and for slowing me down in the downhills where my speed overcomes my endurance. I am forever grateful for your support and never-ending believe in me. Thank you for fostering my ever curious mind and giving me the ability to explore and experience the world. There are no words to express my gratitude.

I want to express my thank you to the funding sources of this thesis. Financial support was provided by the Finnish Cultural Foundation North Savo Regional Fund (grant number 65161550), the Department of Biological and Environmental Sciences of the University of Eastern Finland (PhD student position), Juho Vainio Foundation (grant numbers 201510322, 201610405, 201710136); and NordForsk under the Nordic Programme on Health and Welfare project NordicWelfAir (grant number #75007).
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBR</td>
<td>Brain-to-Body Ratio</td>
</tr>
<tr>
<td>BoD</td>
<td>Burden of Disease</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>DALY</td>
<td>Disability Adjusted Life Years</td>
</tr>
<tr>
<td>DNA</td>
<td>DeoxyriboNucleic Acid</td>
</tr>
<tr>
<td>DOHaD</td>
<td>Developmental Origin of Health And Disease</td>
</tr>
<tr>
<td>DW</td>
<td>Disability Weight</td>
</tr>
<tr>
<td>FGR</td>
<td>Fetal Growth Restriction</td>
</tr>
<tr>
<td>GBD</td>
<td>Global Burden of Disease Study</td>
</tr>
<tr>
<td>HLR</td>
<td>Head-to-Length Ratio</td>
</tr>
<tr>
<td>ICD</td>
<td>International Classification of Disease</td>
</tr>
<tr>
<td>IUGR</td>
<td>IntraUterine Growth Restriction</td>
</tr>
<tr>
<td>LBW</td>
<td>Low Birth Weight</td>
</tr>
<tr>
<td>MBR</td>
<td>Medical Birth Register</td>
</tr>
<tr>
<td>NCD</td>
<td>Non-Communicable Diseases</td>
</tr>
<tr>
<td>OR</td>
<td>Odds Ratio</td>
</tr>
<tr>
<td>PAF</td>
<td>Population Attributable Fraction</td>
</tr>
<tr>
<td>PI</td>
<td>Ponderal Index</td>
</tr>
<tr>
<td>PTB</td>
<td>Preterm Birth</td>
</tr>
<tr>
<td>RNA</td>
<td>RiboNucleic Acid</td>
</tr>
<tr>
<td>RR</td>
<td>Risk Ratio</td>
</tr>
<tr>
<td>SD</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>SE</td>
<td>Standard Error</td>
</tr>
<tr>
<td>SES</td>
<td>SocioEconomic Status</td>
</tr>
<tr>
<td>SGA</td>
<td>Small for Gestational Age</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
<tr>
<td>YLD</td>
<td>Years Lived with Disability</td>
</tr>
<tr>
<td>YLL</td>
<td>Years of Life Lost</td>
</tr>
</tbody>
</table>
DEFINITIONS

Adverse effect is a change in the morphology, physiology, growth, development, reproduction, or life span of an organism, system, or (sub)population that results in an impairment of functional capacity, an impairment of the capacity to compensate for additional stress, or an increase in susceptibility to other influences (IPCS 2004).

Age-weighting in burden of disease estimation incorporates the social value of the time lived at different ages based on the social role at each age, leading to working age being “valued” more than young and elderly age (Murray 1994).

Brain-to-Body Ratio (BBR) is the ratio of brain mass, estimated from head circumference (cm), to body mass (g) and functions as an indicator for body proportionality at birth in this study. A low BBR indicates a small brain in comparison to the body mass (McLennan et al. 1983; Paper III).

\[ BBR = 100 \times \frac{0.037 \times \text{head}^{2.57}}{\text{weight}} \]

Body proportions at birth describe the ratio of weight, length and head circumference to each other. In this work they are expressed as ponderal index, brain-to-body ratio and head-to-length ratio (Paper III).

Burden of Disease (BoD) is a concept to describe loss of healthy life years due to diseases, injuries and risk factors estimated based on the number of years of life a person loses as a consequence of dying early because of a disease (called YLL, or Years of Life Lost); and the number of years of life a person lives with disability caused by the disease (called YLD, or Years lived with Disability). BoD estimates, expressed as DALY (disability adjusted life years) (Murray 1994).

\[ DALY = YLL + YLD \]

Developmental exposure is an exposure during prenatal periods. In this work, maternal smoking is used as an example exposure (Paper IV).

Developmental Origin of Health and Disease (DOHaD), also known as Baker Hypothesis or Fetal Origins of Adult Disease, is a paradigm postulating that exposure to certain environmental influences during critical periods of development and growth may have significant consequences on an individual’s short- and long-term health by developing adaptations that may increase susceptibility to diseases later in life. Long-term, subtle, irreversible changes in the development, structure and function of tissues and vital organs may occur as a result of disruptions in gene expression, cell differentiation and proliferation (Mandy & Nyirenda 2018).
Developmental plasticity is the adaptability to environmental influences leading to a permanent change in phenotype (Hanson & Gluckman 2014).

Disability adjusted life year (DALY) is the unit of burden of disease. It is a year of perfect health lost either due to disease or due to premature death (Murray 1994).

Disability weights (DW) are weights between zero, perfect health, and one, death, used to scale diseases according to the severity in the calculation of years lived with disability (YLD) (Murray 1994).

Discounting in burden of disease estimation is a concept borrowed from economics where the present situation is valued more than the future. As a consequence, the current disease burden associated with a disease is higher now than in the future (Murray 1994).

Epigenetic mechanisms affect gene expression patterns without alterations in DNA base sequence (Hanson & Gluckman 2014).

Hazard is an inherent property of an agent or situation having the potential to cause adverse effects when an organism, system, or (sub)population is exposed to that agent (IPCS, 2004).

Head-to-Length Ratio (HLR) is the ratio of head circumference (cm) to body length (cm) and functions as an indicator for body proportionality at birth in this study. A low HLR indicates a small head in comparison to the body length (Paper III).

Intermediate risk factor within the DOHaD paradigm is a risk factor that is potentially the consequence of a developmental exposure and simultaneously a recognized risk factor for increased disease susceptibility later in life. In epidemiological studies they are commonly endpoints of interest (as adverse birth outcomes) or exposures under study (as marker for suboptimal prenatal development). In this work, low birth weight, preterm birth and childhood overweight are commonly referred to as intermediate risk factors (Paper IV).

Immaturity at birth describes a newborn, who is not fully developed, either due to premature birth or prenatal growth restriction of a combination of both (Hughes et al. 2017).

Intrauterine growth restriction (IUGR) is a pathological small fetal body size due to decreased growth rate diagnosed during prenatal ultrasound scans (Gordijn et al. 2016).
**Intergenerational** effects are induced in one generation passed to subsequent generations via germ cell line (Hanson & Gluckman 2014).

**Life course approach** is based on the DOHaD paradigm and aims at increasing the effectiveness of interventions throughout a person’s life. It focuses on a healthy start to life and targets the needs of people at critical periods throughout their lifetime. It promotes timely investments with a high rate of return for public health and the economy by addressing the causes, not the consequences, of ill health (Jacob et al. 2017).

**Low birth weight (LBW)** is defined as birth weight of less than 2,500 g. In this work it is used as an intermediate risk factor on the pathway from prenatal exposure to later life health effects (Paper III; Paper IV).

**Maternal smoking** refers to active tobacco smoking of pregnant women leading to prenatal exposure of the unborn child (Paper I; Paper II; Paper III; Paper IV; Paper V).

**MATEX (Maternal Exposures) cohort** is a birth cohort identified from the Finnish Medical Birth Register covering practically all births between 1st January 1987 and 31st December 2016 (n=1.7 million mother-child pairs). It was established to study the health effect in the offspring associated with (environmental) exposures during pregnancy.

**Ponderal Index (PI)** is the ratio of body weight (g) to body length (cm) and functions as an indicator for body proportionality at birth in this study. A high PI indicates a high weight in comparison to the body length (Paper III).

\[ PI = 100 \times \frac{weight}{length^3} \]

**Population attributable fraction (PAF)** is defined as the fraction of all cases of a particular disease or other adverse condition in a population that is attributable to a specific exposure (Levin, 1953).

**Preterm birth (PTB)** is a birth before gestational age of 37+0. In this work it is used as an intermediate risk factor on the pathway from prenatal exposure to later life health effects (Paper III; Paper IV).

**Risk** is a function of hazard and exposure. It is the probability of an adverse effect in an organism or (sub)population caused under specified circumstances by exposure to an agent (IPCS 2004).

**Risk assessment** is a process intended to estimate the risk to a given target organism, system or (sub)population, including the identification of attendant uncertainties, following exposure to a particular agent, taking into account the inherent characteristics of the agent of concern as well as the characteristics of the specific target
system. The risk assessment process includes four steps: hazard identification, hazard characterization, exposure assessment, and risk characterization (IPCS 2004).

**Risk characterisation** is an essential part of risk assessment. It is the qualitative and/or quantitative estimation, including attendant uncertainties, of the probability of occurrence or severity of known or potential adverse health effects in a given population. It is often expressed as attributable cases or burden of disease (IPCS 2004).

**Small for gestational age (SGA)** is a small body size for gestational age at birth (Gordijn et al. 2016). In this work it defined separately for weight, body length and head circumference with a cut-off of <10th percentile.

**Teratogenesis** is a process where a *teratogen* causes permanent anatomical or functional disturbances of development, thus having *teratogenic* capacity (Hanson & Gluckman 2014).

**Transgenerational** effects are induced in one generation passed to subsequent generations. The strict definition includes effects not mediated via germ cell line (Hanson & Gluckman 2014).

**Years of Life lived with Disability (YLD)** is the number of years of life a person lives with disability caused by a disease (Murray 1994).

\[
YLD = n \times DW \times L
\]

where n is the number of cases, DW disability weight (0-1) and L duration of disability (in years), which is set to 1 in case of prevalent cases being used.

**Years of Life Lost (YLL)** is the number of years of life a person loses as a consequence of dying early because of a disease (Murray 1994).

\[
YLL = N \times (LE - age_{death})
\]

With N being the number of deaths (for each age and gender), LE the theoretical age and gender-specific life expectancy.
LIST OF ORIGINAL PAPERS

This thesis is based on data presented in the following articles, referred to by the Roman Numerals I-V.


The original publications have been reproduced with permission of the copyright holders.
AUTHOR’S CONTRIBUTION

I. The author designed the study in collaboration with Otto Hänninen, Matti Viluksela, Kirsi Vähäkangas, Mika Gissler, and Heljä-Marja Surcel. The author supported Otto Hänninen in requesting access to data from the MBR. Statistical analyses were done and the first draft of the manuscript was written by the author. All co-authors revised and edited the manuscript.

II. The author designed the study in collaboration with Otto Hänninen, Matti Viluksela, Kirsi Vähäkangas, Mika Gissler and Heljä-Marja Surcel. Statistical analyses were done and the first draft of the manuscript was written by the author. All co-authors revised and edited the manuscript.

III. The author designed the study in collaboration with Otto Hänninen, Matti Viluksela, Mika Gissler and Kirsi Vähäkangas. Statistical analyses were done and the first draft of the manuscript was written by the author. All co-authors revised and edited the manuscript.

IV. The author designed the study in collaboration with Otto Hänninen, Matti Viluksela, and Kirsi Vähäkangas. Literature review and statistical analyses were done by the author. The first draft of the manuscript was written by the author. All co-authors revised and edited the manuscript.

V. The author designed the study in collaboration with Otto Hänninen, Matti Viluksela, and Kirsi Vähäkangas. Literature review and statistical analyses were done by the author. The first draft of the manuscript was written by the author. All co-authors revised and edited the manuscript.
# CONTENTS

1. **INTRODUCTION** .................................................................................................................. 25

2. **REVIEW OF THE LITERATURE** ......................................................................................... 27

2.1 **Burden of disease (BoD)** ............................................................................................... 27

   2.1.1 Burden of disease as extension of risk characterisation .............................................. 27

   2.1.2 Estimation of environmentally attributable disease burden (EBD) ......................... 31

   2.1.3 Burden of Disease in Finland ...................................................................................... 33

2.2 **Developmental origin of health and disease** .................................................................. 34

   2.2.1 Developmental determinants as emerging risk factors for later life disease .......... 34

   2.2.2 Mechanisms linking prenatal exposures and disease susceptibility ..................... 36

   2.2.3 Prenatal development and susceptible time windows of exposure ...................... 38

   2.2.4 Immaturity at birth as a risk factor for later life disease .......................................... 39

2.3 **Maternal smoking** ......................................................................................................... 44

   2.3.1 Composition of tobacco smoke and smoking rates .................................................... 44

   2.3.2 Increasing evidence for health effects in adults and children .................................... 46

   2.3.3 Smoking and the placenta ........................................................................................... 47

3. **AIMS** ................................................................................................................................ 49

4. **MATERIAL AND METHODS** .......................................................................................... 51

4.1 **Study design** .................................................................................................................. 51

   4.1.1 Establishment of the medical birth register-based cohort .......................................... 51

   4.1.2 Exposure assessment .................................................................................................... 51

   4.1.3 Exposure time dependent changes in prenatal growth attributable to maternal smoking ......................................................................................................................................... 52

   4.1.4 Life-long disease burden in the child of maternal smoking ...................................... 54

4.2 **Statistical Analyses** ....................................................................................................... 55

   4.2.1 Exposure analyses ........................................................................................................ 55

   4.2.2 Epidemiological analyses .............................................................................................. 56

   4.2.3 Chained burden of disease model ................................................................................ 56

   4.2.4 Meta-Analysis of childhood cancer studies .................................................................. 57

4.3 **Ethics approval and register data permit** ...................................................................... 58

5. **RESULTS** .......................................................................................................................... 59

5.1 **Register-based approach for identification early markers of effect and life-long follow up** ........................................................................................................................................................................ 59

   5.1.1 Baseline MATEX cohort for life-long follow up ......................................................... 59

   5.1.2 Exposure characteristics .............................................................................................. 61

   5.1.3 Growth restriction attributable changes in body proportions at birth .................... 63

5.2 **Life-long disease burden of developmental origin** ....................................................... 66

   5.2.1 Health effects of maternal smoking and intermediate risk factors .......................... 66

   5.2.2 Evidence for an association of maternal smoking with childhood cancer ............ 69

   5.2.3 Burden of disease in the child attributable to maternal smoking and intermediate risk factors ........................................................................................................................................ 70

6. **DISCUSSION** .................................................................................................................... 73

6.1 **Register-based epidemiology** ....................................................................................... 74
6.1.1 Register linkage for life-long follow up ............................................................... 74
6.1.2 Mechanistic interpretations ............................................................................... 76
6.1.3 Uncertainties of register-based analyses ......................................................... 78

6.2 Burden of disease of developmental origin .......................................................... 80
   6.2.1 Potential of DOHaD to explain disease burden ............................................. 80
   6.2.2 Health endpoints in the causal chain ............................................................ 83
   6.2.3 Uncertainties in the chained risk model ....................................................... 84

6.3 Conceptual uncertainties ....................................................................................... 85

6.4 Future perspectives ............................................................................................... 86

7 CONCLUSIONS ....................................................................................................... 89

8 REFERENCES .......................................................................................................... 91
1 INTRODUCTION

With improving health and treatment options, non-communicable diseases (NCD) have become more important for both health professionals and the general public. The latter are almost bombarded by recommendations for healthy lifestyle and how to reduce risks for common chronic diseases, such as no smoking, nor alcohol to prevent cancer and exercise and healthy diet to prevent cardiovascular diseases. Although it may seem like the most important risk factors have been identified, a major share of chronic diseases remains unexplained at this point. In Finland, about 1.4 million healthy life years were lost due to NCD in 2017. The Global Burden of Disease (GBD) study estimates that about 44% of the disease burden can be explained by 84 well-established risk factors, such us unhealthy diet, smoking, occupational exposures and air pollution, leaving 56% of disease burden unexplained (Stanaway et al. 2018). Thus, disease prevention is seriously limited by lack of knowledge on underlying potentially modifiable risk factors.

The Developmental Origin of Health and Disease (DOHaD) paradigm proposes the environment during pre- and early postnatal development as a determinant for later life susceptibility for disease later in life (Barker et al. 2007). Epigenetic changes, which are heritable changes of phenotype not due to change in the DNA sequence but due to altered gene activity and expression, are understood to be one of the responsible mechanisms (Berger et al. 2009). The prenatal period is of special concern since the organogenesis is a highly organized period of cell proliferation, migration and differentiation. Additionally, the period of gametogenesis and fertilization is marked by significant epigenetic reprogramming consisting of a general demethylation followed by re-methylation. Disturbance during the reprogramming can trigger epigenetic imbalances and modify the susceptibility for disease development (Alvarado-Cruz et al. 2018). The DOHaD paradigm was built upon the observation of correlation between areas of high infant mortality with high incidence of ischaemic heart disease in the adult population (Barker et al. 2007), and the observation that sub-optimal maternal nutrition was associated with higher risk for cardiovascular disease and diabetes in the adult child (Roseboom et al. 2006). Soon, the paradigm was widened to include maternal lifestyle, diseases and other stressors that may interfere with the fetal development. It is suggested that DOHaD and epigenetic mechanisms have the potential to shed light into the etiology of complex chronic diseases and with that help to identify modifiable risk factors (Heindel & Vandenberg 2015).

To study the potential of developmental determinants as a risk factor for later life disease, as proposed by DOHaD, two requirements need to be met: Firstly, the suspected risk factor should be reasonably common in order to be able to detect associations with later life diseases in epidemiological studies. Ideally, it should be modifiable, so that it can be used for public health interventions. Secondly, a well-sized birth cohort is needed to follow up the child into advanced adulthood, since most chronic disease only emerge in the second half of life.
In this study, maternal smoking was selected as a readily modifiable risk factor, which still affects up to 15% of all unborn children in Finland. The aim of this work was to explore maternal smoking as a developmental determinant of health and disease throughout the life course by analysing the effect of maternal smoking on growth restriction associated changes in body proportions and estimating the long-term disease burden of maternal smoking and associated risk mediators in a chained risk model.
2 REVIEW OF THE LITERATURE

2.1 BURDEN OF DISEASE (BOD)

2.1.1 Burden of disease as extension of risk characterisation

Public health protection has a high need for priority setting, since limited resources should be spent effectively, efficiently and fairly. Decision makers who allocate funding need to decide between the relative importance of diseases and conditions (Murray 1994). The relative effectiveness of treatments and interventions is hard to compare, since it is intrinsically difficult to compare different diseases. Interventions and risk management options in environmental health work need to be comparable on the level of achievable improvements in public health. As a solution, composite indices of health combining morbidity and mortality, such as disability adjusted life years (DALY), have been proposed (Murray 1994).

The concept of an index of health has already been discussed in the 1960s (Sullivan 1966). The suggested approaches range from a “functional adequacy” measure (Sanders 1964) to probabilistic modelling of weighted disease characteristics (Chiang 1965). However, the approaches were deemed not fit for the purpose of a national health index (Sullivan 1966). An index combining morbidity and mortality was subsequently published by Sullivan (1971). The concept of an index of health became widely popular with the development of Burden of Disease (BoD) as input into the World Bank’s World Development Report for 1993, later called Global Burden of Disease (GBD) study (Murray 1994; Anand & Hanson 1997). There was a need for an index that would guide setting priorities for health service and health research, support identification of disadvantaged groups and targeting health interventions and provide comparable measure of output for intervention, programme and health sector evaluation and planning (Murray 1994). BoD was aimed to combine mortality and morbidity into assessments of health, to produce objective disease burden assessment, and lastly to aid in cost-effectiveness assessment of interventions in terms of cost per unit of disease burden averted (Murray & Lopez 1996).

BoD was developed taking age and gender into account. Country-specific life expectancy and household income were not taken into account since it would bias BoD estimates towards countries with higher life expectancy, since more years can potentially be lost due to premature death (Murray 1994).

BoD estimates, expressed as DALY, are the sum of morbidity (Years Lived with Disability, YLD) and a mortality (Years of Life Lost, YLL) component (Equation 1). The combination of morbidity and mortality is enabled by scaling morbidity in comparison to mortality using disability weights (DW). The DW range from zero, perfect health, to one, death (Murray 1994).
\[ \text{DALY} = \text{YLL} + \text{YLD} \]

YLL and YLD are defined according to Hänninen and Knol (2011) as following (Equation 2, Equation 3):

\[ \text{YLL} = N \times (L - \text{age}_\text{death}) \]

with \( N \) being the number of deaths (for each age and gender), \( L \) the theoretical age and gender-specific life expectancy, and \( \text{age}_\text{death} \) the age at death.

\[ \text{YLD} = n \times \text{DW} \times L \]

where \( n \) is the number of cases, \( \text{DW} \) disability weight (0-1) and \( L \) duration of disability (in years), which is set to 1 in case of prevalent cases being used.

The Global Burden of Disease (GBD) project was developed from the first global BoD assessment efforts for the World Development Report (Murray 1994). Since its first publication, the GBD methods and estimates have been updated several times and the lead changed from the WHO and World Bank to the Institute for Health Metrics and Evaluation (Table 1). The updates included widening the scope of the project as well as updating the methods. The method updates reflect changing views on value choices, such as age-weighting and discounting, as well as updated input data, e.g. life expectancy and shift from incidence based estimations to prevalence based ones. DW have been derived by comparing diseases with each other and sorting them according to the severity of the effect. Originally, this work was done solely by public health experts, while in later updates the general public was consulted (Salomon et al. 2013; Salomon et al. 2015).

The BoD concept has been criticised as inherently flawed due to the lack of a clear definition of “health” and the ignorance of several dimensions of health, which are not commensurable, making the relation “healthier than” incomplete (Hausman 2012). Solberg et al. (2018) argue that the claimed equivalence between YLD and YLL is untrue and the components are actually incommensurable, meaning they cannot be compared or summed. According to them, the incommensurability arises from the opinion that YLL; meaning death, presents an actual burden. Their argumentation is based on the understanding of the DALY as an aggregated individual burden that aims to measure health impairment directly (Solberg et al. 2018). It has further been discussed whether the BoD should be measured by its consequences on health but not well-being (Hausman 2012; Broome et al. 2002). In addition, BoD measured in DALY in cost-effectiveness analyses has been discussed to ignore non-health sector returns (e.g. economic returns), providing an incomplete picture (Anand & Hanson 1997).

The detailed reporting of the methods and input data used in each published study is spread across several publications and online supplemental materials. Complications include e.g. sequelae handling in the GBD2004 update (Schroeder 2012) and the difficulties tracing the life expectancy used in the GBD2017 update. No clear life
expectancy at birth applied to the YLL estimation seems to be reported. This hinders the comparison of the extent of trends in YLL are driven by actual changes in mortality and by how far they are driven by increasing theoretical life expectancy.

Despite the aim of objectiveness, value choices impact the BoD estimates beyond the choice of DW; such as age-weighting, discounting, and life expectancy. The early BoD estimates were age-weighted and discounted (Table 1). Age-weighting incorporates the social value of the time lived at different ages based on the social role at that age, leading to working age being “valued” more than young and elderly age. Discounting is a concept borrowed from economics where present situation is valued more than future. Age-weighting and discounting result in one disease case causing more burden in the present than in future and in working age group than in children or the elderly (Murray 1994). Life-expectancy was chosen based on the healthiest known population (Japan for GBD 1990, South Korea for 2001) and later life table modelling (Table 1). After a debate on the usefulness of age-weighting and discounting (Voight & King 2014; Arnesen & Kapingiri 2004; Anand & Hanson 1997), they were not applied in the GBD2010 update anymore (Table 1). Questions have been raised about the purpose of DW: quantity of health, value of health state or well-being (Schroeder 2012). Additionally, it was argued that it is impossible to generalise DW across all socioeconomic groups, social backgrounds and general health states. Furthermore, it was criticised that health cannot be separated from welfare, which is the result of disease symptoms and their interaction with the environment (Voight & King 2014). Another point of criticism has been the inconsistency in health state definitions. For some endpoints co-morbidities have been included in GBD2004, while it was reported that it was never done (Schroeder 2012). Additionally, in GBD2010 social implications have been taken into account for some health states, while for others this was not the case (Voight & King 2014). It was acknowledged that DW are sensitive to these inconsistencies (Salomon et al. 2012). It even has been argued that the YLD should be estimated without a DW, solely based on incidence and duration with a description of the condition in question to avoid the big impact of subjectiveness introduced by the DW, on the BoD (Arnesen & Kapingiri 2004).
Table 1. Methodological summary of key Burden of Disease Studies.

<table>
<thead>
<tr>
<th></th>
<th>GBD1990</th>
<th>GBD2004</th>
<th>GBD2010</th>
<th>GHE2012</th>
<th>GBD2017</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>the BoD estimation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Updates</strong></td>
<td></td>
<td>2004; 2009</td>
<td>Annually</td>
<td>2016</td>
<td>na</td>
</tr>
<tr>
<td><strong>Diseases</strong></td>
<td>131</td>
<td>159</td>
<td>241</td>
<td>241</td>
<td>359</td>
</tr>
<tr>
<td><strong>Sequelae</strong></td>
<td>483</td>
<td>474</td>
<td>1,160</td>
<td>1,160</td>
<td>1,410</td>
</tr>
<tr>
<td><strong>Risk factors</strong></td>
<td>10</td>
<td>26</td>
<td>69</td>
<td></td>
<td>84</td>
</tr>
<tr>
<td><strong>Age groups</strong></td>
<td>5</td>
<td>8</td>
<td>21</td>
<td>7</td>
<td>21</td>
</tr>
<tr>
<td><strong>Regions</strong></td>
<td>8</td>
<td>14</td>
<td>20</td>
<td>Not reported</td>
<td>21</td>
</tr>
<tr>
<td><strong>Life expectancy</strong></td>
<td>Females 82.5 years</td>
<td>Males 80 years</td>
<td>86 years for both sexes</td>
<td>92 years for both sexes</td>
<td>Not reported</td>
</tr>
<tr>
<td><strong>used in YLL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>estimation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>YLD</strong></td>
<td>Incidence (prevalence based also calculated)</td>
<td>Incidence</td>
<td>Prevalence</td>
<td>Prevalence</td>
<td>Prevalence</td>
</tr>
<tr>
<td><strong>approach</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Discounting of future burden</strong></td>
<td>3% (0% also calculated)</td>
<td>3%</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td><strong>Age-weighting</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td><strong>Disability weights</strong></td>
<td>Health expert based</td>
<td>Same as GBD1999 with ad hoc assignment for newly added diseases</td>
<td>Disability weights based on health experts and general public Adjusted for comorbidities</td>
<td>GBD2010 values with revised lay descriptions</td>
<td>Disability weights based on health experts and general public Adjusted for comorbidities</td>
</tr>
<tr>
<td><strong>Co-morbidities</strong></td>
<td>Excluded</td>
<td>Excluded</td>
<td>Calculated</td>
<td>Calculated</td>
<td>Calculated</td>
</tr>
<tr>
<td><strong>Collaborators</strong></td>
<td>World Bank, WHO, Harvard School of Public Health</td>
<td>World Bank, WHO, Harvard School of Public Health</td>
<td>IHME, University of Queensland, Harvard School of Public Health, Johns Hopkins Bloomberg School of Public Health, University of Tokyo, Imperial College London, WHO</td>
<td>WHO</td>
<td>IHME, University of Queensland, Harvard School of Public Health, Johns Hopkins Bloomberg School of Public Health, University of Tokyo, Imperial College London, WHO</td>
</tr>
</tbody>
</table>

GBD: Global burden of disease study; GHE: Global Health estimates; YLL: Years of life lost; YLD: Years lived with disability; na: not applicable; IHME Institute for Health Metrics and Evaluation
Despite of the various criticisms, BoD is currently a widely used health index that accounts for mortality and morbidity. It is routinely used as a metric in risk characterisation during health risk assessment efforts. Environmental exposures and other health risks are seldom associated with only one health outcome and therefore the pooling of all possible loss of health is crucial. This summary of health loss is then subsequently used in risk management, where different mitigation and intervention options are evaluated based on effectiveness and efficiency (Figure 1). The composite nature of BoD enables the characterisation and evaluation as a summary, as well as by its separate components, morbidity and mortality, or for population subgroups (Murray 1994).

2.1.2 Estimation of environmentally attributable disease burden (EBD)

Understanding of the contributions of risk factors to the disease burden is essential for prioritisation of interventions and effective disease prevention. The population attributable fraction (PAF) is used to calculate the proportion of the outcome that would theoretically not have occurred without exposure to the risk factor. The idea of PAF estimation was first introduced by Levin (1953).

It was introduced as shown in Equation 4 with \( P(E_2) \) as exposure prevalence within the entire population, \( RR = \frac{R_1}{R_2} \) as the relative risk between exposed and unexposed individuals, \( R_1 = P(D_2|E_1) \) as outcome occurrence in the unexposed and \( R_2 = P(D_2|E_2) \) outcome occurrence in the exposed.

\[
PAF = \frac{P(E_2)(RR-1)}{1+ P(E_2)(RR-1)} \tag{4}
\]

In general, a difference is made between attributable fraction and PAF. Attributable fraction is focused on those individuals with the outcome. It estimates the proportion of
the outcomes occurring among the exposed individuals, which is in excess in comparison to the unexposed individuals (= the proportion of the outcome among the exposed individuals attributable to the given exposure). It restricts the attention to exposed cases and depends only on the strength of association (RR). In contrast, the population attributable fraction (PAF) depends both on the strength of the association and the prevalence of the exposure, in order to quantify the importance of an exposure at population level by shifting the focus onto the entire population. Therefore, it generalises the attributable fraction to the total population of exposed and unexposed individuals.

Several names have been used for attributable fraction and PAF. Attributable fraction has also been referred to as attributable risk (MacMahon & Pugh 1970) and attributable risk percent (Cole & MacMahon 1971). Names used for PAF include attributable risk (Walter 1975), population attributable risk (MacMahon & Pugh 1970), attributable proportion (Levin 1953), population attributable fraction (Deubner et al. 1975), etiologic fraction (Miettinen 1974), excess fraction (Greenland & Robins 1988), attributable risk percentage (Sturmans et al. 1977), and population attributable risk percent (Cole & MacMahon 1971). It is worth noting, however, that the same names may not have a similar mathematical definition. As an example, Miettinen (1974) used attributable fraction and etiological fraction synonymously to describe fraction of disease that would have not occurred had the factor been absent from the population. In contrast, Greenland and Robins (1988) define etiological cases as those cases, where the risk factor contributed to the development of outcome, but may have not been totally prevented by the absence of exposure. They use excess cases for the fraction, which theoretically would be preventable by the absence of exposure. The lack of consensus on terms and definitions requires a clear definition of used terms in publications.

Different estimation methods have been developed to take into account confounding factors (Walter 1976; Bruzzi et al. 1985; Benichou 2001), sequential chain of effects (Eide & Gefeller 1995; Mason & Tu 2008) and specific study methods, e.g. cohort studies with censored time-to-event data (Chen et al. 2006; Samuelsen & Eide 2008; Cox et al. 2009), total mortality in cohort studies (Laaksonen, Knekt et al. 2010) and incidence in cohort studies with censoring due to death (Laaksonen, Härkänen et al. 2010).

A systematic assessment of 84 established risk factors, ranging from environmental and occupational exposures to nutrition and lifestyle factors, concluded that roughly 48% (1.2 billion DALY) of the global disease burden in 2017 could be explained by those risk factors. The risk factor attributable fraction of NCD burden was estimated to be 45.6% (Stanaway et al. 2018). In Europe, smoking, alcohol and metabolic risk factors (high systolic blood pressure, high BMI) are the leading risk factors for disease burden. Despite systematic efforts to identify and quantify risk factors for disease development, roughly half of disease burden remained unexplained by established risk factors.
2.1.3 Burden of Disease in Finland

In Finland in 2017 the disease burden was dominated by NCD (86% of total BoD; 1,384,091 DALY), with only minor parts attributable to injuries (10% of the total BoD; 160,143 DALY) and communicable, maternal, neonatal and nutritional diseases (4% of total BoD; 61,033 DALY) (GBD 2018 Results tool). The total disease burden was higher in men than in women, driven by higher burden in NCD and injuries. Similarly, the fraction of BoD explainable by 84 established risk factors was higher in men than in women (GBD 2018 Results tool). Overall the disease burden increases with increasing age (Figure 2).

Figure 2. Burden of Disease in Finland in 2017 (A) attributable and unexplained fraction by metrics and gender; (B) age dependence of BoD. Data from GBD 2017 (2018) results tool.
Since 1990 the relative importance of NCD increased by roughly 7% in Finland, equalling about 56,000 DALY more in 2017 than in 1990, which was partly due to aging population, better treatment options and prevention of infectious diseases. The five leading causes of disease burden remained unchanged between 1990 and 2017, namely: cardiovascular diseases, neoplasms, musculoskeletal disorders, neurological disorders and mental disorders. Diabetes and chronic kidney disease increased in their relative importance for the total BoD. Ischaemic heart disease and stroke, low back pain, Alzheimer’s disease, diabetes, depressive and anxiety disorders, lung and breast cancer, and chronic obstructive pulmonary disease were the important diseases in terms of lost healthy life. Neonatal disorders and especially lower respiratory infections were associated with lower relative disease burden in 2017 compared to 1990 (GBD 2018 Compare).

Considerable work has been reported about environmentally attributable disease burden in Finland. Studied environmental risk factors range from air pollution, noise, methylmercury, dioxins and radon to second hand tobacco smoke (Hänninen et al. 2014; Lehtomäki et al. 2018; Asikainen et al. 2013). In 2015, roughly 35,000 DALY were attributed to air pollution in Finland (Lehtomäki et al. 2018). Furthermore, roughly 5,000 DALY were attributable to radon, and another 5,000 DALY to second hand tobacco smoke and 1,700 DALY to dioxins (Hänninen & Knol 2011). According to the systematic assessment of risk factors in the GBD2017 study, roughly 44% of the NCD burden was attributable to established risk factors. Thus, currently the majority of disease burden in Finland cannot be attributed to known risk factors, limiting the options for public health interventions.

2.2 DEVELOPMENTAL ORIGIN OF HEALTH AND DISEASE

2.2.1 Developmental determinants as emerging risk factors for later life disease

The Developmental Origin of Health and Disease (DOHaD) paradigm can at least partly contribute to better understanding of yet unidentified risk factors for disease development. It was long assumed that chronic diseases are a consequence of genetic predisposition, partly mediated via single nucleotide polymorphisms, and voluntary lifestyle choices, such as smoking, diet and physical activity (Hanson & Gluckman 2014). Public health interventions focused on these lifestyle choices had limited success in disease prevention (Hanson & Gluckman 2014; Barker 2007). Emerging evidence is indicating that genotype and phenotype do not have a static relationship, but that one genotype can give rise to a variety of phenotypes depending on developmental and environmental processes (Hanson & Gluckman 2014). This concept of developmental plasticity is a cornerstone of the DOHaD paradigm, which suggests that exposures during sensitive periods of development can alter disease susceptibility in later life. The underlying mechanisms are proposed to be disruption of cell differentiation, cell
proliferation and altered gene expression leading to, in some cases subtle, irreversible changes in the development, structure, and function of tissues and organ systems (Mandy & Nyirenda 2018). While the DOHaD paradigm is used to explain disease susceptibility and diseases itself, some of the underlying mechanisms are also important for healthy responses to challenges during life. The paradigm includes the notion that phenotype changes due to developmental environment and altered responses to challenges can both be within the normal physiological range. This is the essence how adaptation to various changes in environment can happen with no or minimal obvious adverse effects although the genotype does not change (Hanson & Gluckman 2014).

Comparable ideas as the DOHaD paradigm have been also published under the terms Barker Hypothesis, Fetal Origins of Adult Disease Hypothesis and Fetal Programming Hypothesis, and Advanced Fetal Programming Hypothesis (Reichetzeder et al. 2016).

Earliest examples of the long-term health effects of exposures during prenatal development include famine, industrial accidents and medication use during pregnancy. The importance of developmental exposures on disease susceptibility was widely recognized only after retrospective studies of famine and maternal undernutrition across Europe consistently showed a strong association between maternal undernutrition, low birth weight and later cardiovascular disease (Roseboom et al. 2006; Baker 2007). Industrial disasters, such as the mercury contamination in Minamata Bay, Japan, and polychlorinated biphenyls contamination of rice oil in Japan, led to widespread exposure among pregnant women and subsequent neurodevelopmental problems (Heindel et al. 2017). Environmental exposures to chemicals and lifestyle factors are increasingly studied in the context of DOHaD since the end of the 1980s and early 1990s focusing on exposure to polychlorinated biphenyls (Gladen et al. 1988; Gladen et al. 1990; Jacobsen et al. 1990). Shortly afterwards, the effects of prenatal exposure to diethylstilboestrol and pesticides were reported (Waggoner et al. 1994; Meinert et al. 1996). Most studies focus common chemical exposures with well-established effects after exposures in adults, such as polychlorinated biphenyls, pesticides, polyaromatic hydrocarbons and methylmercury (Heindel et al. 2017). Regularly studied health endpoints include problems in organ systems that are known to be especially sensitive to insults during prenatal development, such as central nervous system and the respiratory system. Additionally regularly studied endpoints include diseases of great public health concern, such as cancer and metabolic impairments, especially obesity and diabetes (Heindel et al. 2017). Heindel and colleagues (2017) point out that 90% of studies of developmental exposures to environmental chemicals investigated health endpoints during childhood with only a few studies on adult health focusing on breast and testicular cancer. They suggest that studies with longer follow up time are needed. Most endpoints included in published studies are part of the etiological chain of various chronic diseases, and they affect several tissues and pathways (Heindel et al. 2017).
The scope of the DOHaD paradigm is continuously growing with new evidence. In contrast to previous proposals that low birth weight is the first symptom in the causal pathway from fetal origin to chronic diseases, it is now thought that long-term effects of fetal origin can be independent of birth weight (Hanson & Gluckman 2014) or at least associated in a U-shaped manner with disease susceptibility increasing towards the extreme ends of the range of birth weight (Reichetzeder et al. 2016; Wadhwa et al. 2009). The importance of growth in early postnatal life has been identified as an important marker of adult disease risk (Reichetzeder et al. 2016; Wadhwa et al. 2009). Moreover, the effect of paternal nutrition and health on the child’s disease susceptibility is emerging (Reichetzeder et al. 2016). The importance of inheritance of epigenetic changes for developmental origin of diseases is discussed (Reichetzeder et al. 2016; Hanson & Gluckman 2014). Epigenetic modifications have the potential to be intergenerationally transmitted via the germ cell line. The potential of intergenerational effects of the paternal line mediated via epigenetic changes in sperm has not been well studied so far and clear understanding of the magnitude of effect is missing (Soubry 2018). The maternal germ cell line is potentially more susceptible for intergenerational transmission of effect than the paternal line based on the development of primordial ovarian follicles during fetal development. Thereby maternal genetic information, epigenetically modified during prenatal development, can be passed to the grandchild at the time of fertilization (Hanson & Gluckman 2014).

The relevance of the DOHaD paradigm has been criticised. The validity of low birth weight (LBW) as an indicator for growth restriction has been questioned and thereby the reliability of study results relying on LBW as indicator for prenatal development (Simmons 2005). Furthermore, the design of DOHaD studies was criticised as being unable to account for unmeasured confounding from genetics and familial factors (Phillips 2006). These study design issues contribute to the conflicting results and small effect sizes observed in DOHaD studies. It was argued that publication bias leads to the inflation of effect sizes (Phillips 2006).

2.2.2 Mechanisms linking prenatal exposures and disease susceptibility

The idea that disruptions during developmental periods and developmental plasticity affect susceptibility for later life diseases is nowadays widely accepted, while some level of disagreement about the mechanisms remain (Reichetzeder et al. 2016). Often, the focus has been on the effect of nutrition during developmental periods while neglecting other exposures during the same time. The Thrifty Phenotype hypothesis or Mismatch Concept, for example, suggest a mismatch of nutrient availability prenatally and postnatally and associated programming that is not beneficial for postnatal nutrient richness (Calkins & Devaskar 2011). In contrast, the Fetal Insulin Hypothesis postulates that the genotype is the main contributor to later life diabetes risk. According to this theory, low birth weight is the first result of impaired insulin secretion or sensitivity (Reichetzeder et al. 2016; Calkins & Devaskar 2011). Recently, Li et al. (2017) proposed several mechanisms on three levels that contribute to later life disease susceptibility:
(epi)genetic factors, detrimental factors during developmental periods, and environmental factors later in life. While mechanisms at each level can cause disease independently, the cooperation of mechanisms leads to increased risk and highlights the gene-environment interaction in DOHaD (Li et al. 2017).

Traditionally, subtle changes in later life disease susceptibility were not an endpoint of concern in developmental toxicity. Understandably, the main concerns were fetal death and stillbirth, as well as congenital anomalies and growth restriction, all evident soon after birth (NRC 2000). It was widely recognized that the susceptibility to teratogens depends on the genotype of the mother and the child (NRC 2000). Growth restriction has been understood as the result of high level of widespread cell death that cannot be replenished by available repair and compensatory mechanisms. If extensive cell death occurred in essential organ systems, prenatal death was assumed to be the consequence (NRC 2000).

Generally, the mechanisms of pathogenesis include the effects on gene expression, patterns of apoptosis, replication, cell cycle, cell proliferation, migration, adhesion, secretion, endocytosis, uptake, and signal transduction (NRC 2000). Additionally, exposures during prenatal development may interfere with receptor-ligand interactions by covalent binding, peroxidation of lipids and proteins, interference with sulfhydryl groups, inhibition of protein function and maternally mediated effects (NRC 2000).

Developmental programming is generally understood to be mediated via epigenetic changes, changes in cell-cycle regulation, and changes in cellular or tissue differentiation. In animal studies, for example, maternal malnutrition was associated with increased apoptosis in pancreatic islets and shortened telomere length, indicating accelerated cell cycling in the offspring. Furthermore, reduced uteroplacental perfusion was associated with reduced neuronal count in some brain regions. There are many well-described examples of altered tissue differentiation in animal models after maternal malnutrition, for example in blood vessels, kidney, and muscles. Also, altered enzyme expression pattern have been described (Gluckman & Hanson 2004).

Some of the important molecular mechanisms thought to be behind developmental programming are epigenetic mechanisms. Epigenetics deals with the heritable change of phenotype not due to changes in the DNA sequence but due to altered gene activity and expression (Hanson & Gluckman 2014). Epigenetic modifications that alter the frequency of gene expression include changes in DNA methylation, histone modifications or non-coding RNA (Wadhwa et al. 2009; Reichetzeder et al. 2016; Dolinoy et al. 2007; Schug et al. 2013). The term was first defined in the 1940s as the interactions of genes with their environment, which bring the phenotype into being without a clear understanding how this interaction works at biomolecular level. Three decades later it was proposed that the chemical, covalent DNA modifications, including DNA methylation, are the molecular mechanism behind the hypothesis (Dolinoy et al. 2007). Epigenetic alterations and thereby abnormal expression of imprinted genes have been identified as important contributory, if not causative, factors in the development of syndromes such as Prader-Willi, Angelman and Beckwith-Wiedermann (Dolinoy et al. 2007).
Epigenetic modifications alter gene expression by changes in accessibility of DNA to transcription factors, the efficiency of transcription and stability of transcribed mRNA (Reichetzeder et al. 2016). DNA methylation, the addition of a methyl group at the C5 position of the cytosine pyrimidine ring, alters the access of transcription factors leading to a decreased transcription frequency by attracting methyl binding proteins that trigger a more condensed chromatin packing (Hanson & Gluckman 2014; Reichetzeder et al. 2016). These changes can be stable, leading to a permanent change in methylation of promoter regions of genes thereby altering gene expression in long-term (Reichetzeder et al. 2016). Chromatin packing is the coiling of DNA around octamers of histone proteins (Dolinoy et al. 2007). Roughly, the looser the chromatin packing the more a gene is expressed. Histone proteins can be modified which in turn alters the tightness of chromatin packing. The effect of histone methylation depends on the exact spot of methylation, while histone acetylation generally loosens the chromatin packing and increases transcription (Dolinoy et al. 2007). The consequences of non-coding RNA modification are various and include gene suppression and as well as activation (Chen et al. 2017). Non-coding RNA can trigger RNAse activity resulting in elimination of mRNA transcribed by target genes (Reichetzeder et al. 2016). Epigenetic changes may be tissue specific, or at least are thought to have tissue specific effects (Wadhwa et al. 2009).

The mechanisms and timing of exposure mitigate a wide range of effects from severe malformations and mortality to subtle changes in organ function across all organ systems. This complicates epidemiological analysis of the long-term effects of prenatal exposures since a multitude of subtle changes in many organ systems need to be studied.

2.2.3 Prenatal development and susceptible time windows of exposure

Sensitive windows of development dictate the potential effects of developmental exposures (Wadhwa et al. 2009; Alvarado-Cruz et al. 2018). Originally, the degree of sensitivity to developmental toxicant-induced structural malformations was understood to be low during the pre-implantation phase, maximal during organogenesis, and low during fetal development (NRC 2000). In general, early pregnancy has been shown to be most sensitive to adverse exposures. Each organ system differs in timing of the most sensitive window (Figure 3). More subtle changes, such as functional and susceptibility, may extend beyond the sensitive time windows for major congenital malformations. Germ cells have been recognized as sensitive to exposures within the DOHaD framework (Soubry 2018).

Prenatal development is not only characterized by highly coordinated cell proliferation, differentiation, migration and apoptosis, but also by global epigenetic changes. During gametogenesis a general demethylation is followed by a remethylation in primordial germ cells. During fertilization both parental genomes undergo epigenetic modifications starting with active demethylation of the paternal DNA, followed by passive demethylation of the maternal genes. This is thought to restore the totipotency
of the fertilized egg (Dolinoy et al. 2007; Reichetzeder et al. 2016; Alvarado-Cruz et al. 2018). Epigenetic changes may still occur well after birth to some degree (Hanson & Gluckman et al. 2014; Reichetzeder et al. 2016; Schug et al. 2013).

Figure 3. Sensitive windows of organ development in human. Modified from Moore et al. 2013.

Especially during early pregnancy, when the women are not aware of the pregnancy, lifestyle choices may not be modified yet to accommodate for the sensitivity of a fetus. Alcohol consumption, smoking, occupational exposures and medication usage may occur. The effects of any exposure potentially differ greatly by time of exposure during pregnancy. Roseboom et al. (2006), for example, showed the exposure time-dependent long-term effects of undernutrition in the Dutch Hunger Winter cohort.

2.2.4 Immaturity at birth as a risk factor for later life disease

Immaturity has long been recognised as potentially the most important risk factor for morbidity and mortality in early life. LBW was used as an indicator for sub-optimal prenatal development in the earlier DOHaD studies when no precise information on maternal nutrition or prenatal growth was available. Prematurity was used to describe any newborn that was less physically developed than expected. In 1923, for example, prematurity was proposed to be assessed based on weight, length, and general characteristics, such as texture of the skin, underdeveloped nails, crying, unstable temperature and history of expected birth (Hughes et al. 2017). In the International Classification of Diseases (ICD) underdevelopment at birth is included under the summary of “Disorders related to length of gestation and fetal growth” (Table 2) (WHO 2004).
Table 2. Disorders related to length of gestation and fetal growth (modified from ICD-10 (WHO 2004)).

<table>
<thead>
<tr>
<th>Code</th>
<th>Diagnoses</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>P05</td>
<td>Slow fetal growth and fetal malnutrition</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P05.0 Light for gestational age</td>
<td>Usually refers to weight below but length above 10th percentile for gestational age</td>
</tr>
<tr>
<td></td>
<td>P05.1 Small for gestational age</td>
<td>Usually refers to weight and length below 10th percentile for gestational age</td>
</tr>
<tr>
<td></td>
<td>P05.2 Fetal malnutrition without mention of light or small for gestational age</td>
<td>Infant, not light or small for gestational age, showing signs of fetal malnutrition, such as dry, peeling skin and loss of subcutaneous tissue</td>
</tr>
<tr>
<td></td>
<td>P05.9 Slow fetal growth, unspecified</td>
<td>Fetal growth retardation NOS</td>
</tr>
<tr>
<td>P07</td>
<td>Disorders related to short gestation and low birth weight, not elsewhere specified</td>
<td>Note: When both birth weight and gestational age are available, priority of assignment should be given to birth weight</td>
</tr>
<tr>
<td></td>
<td>P07.0 Extremely low birth weight</td>
<td>Birth weight of 999 g or less</td>
</tr>
<tr>
<td></td>
<td>P07.1 Other low birth weight</td>
<td>Birth weight of 1,000-2,499 g</td>
</tr>
<tr>
<td></td>
<td>P07.2 Extreme immaturity</td>
<td>Less than 28 completed weeks (less than 196 completed days) of gestation</td>
</tr>
<tr>
<td></td>
<td>P07.3 Other preterm infants</td>
<td>28 completed weeks or more but less than 37 completed weeks (196 completed days but less than 259 completes days) of gestation</td>
</tr>
</tbody>
</table>

While diagnostic codes under P05 take gestational age into account, diagnostic codes P07.0 and P07.1 assess birth weight in isolation. Birth weight and gestational age are highly correlated thereby not all children born with LBW are necessarily small or light for gestational age (SGA). A first demarcation between LBW and preterm birth (PTB) was proposed by Capper in 1928. He noted that more than 70% of LBW newborns show no other signs of prematurity besides their birth weight (Hughes et al. 2017). The cut-off value for both, LBW and PTB are arbitrarily chosen (Hughes et al. 2017; Chawanpaiboon et al. 2019; Engle 2006). Even the current markers for immaturity are rather crude and cannot fully distinguish healthy newborns from those, who are small due to suboptimal development. Investigation of possible long time prognosis of health is complicated by the grouping of different underlying causes into the phenotype.

In clinical context, growth restriction is routinely divided into symmetrical and asymmetrical growth restriction. Symmetrical growth restriction in general is associated with poorer prognoses for long-term health (Sharma, Shastri et al. 2016). In symmetrical growth restriction, all anthropometric measures are reduced: weight, length and had circumference in newborns or abdominal circumference, bipartial diameter, femur length and head circumference in the fetus, while the reduction of weight (or abdominal circumference) is more pronounced in asymmetrical growth restriction (Sharma, Shastri et al. 2016). This division ignores the other end of asymmetry: newborns with average weight but reduced length and head circumference. Asymmetrical growth restriction is also called “Wasting” (Victoria et al. 2015). The etiology of this type of growth restriction
often includes acute adverse exposure in the last weeks of pregnancy when fat accumulation takes place. A reduction in body length is called “Stunting” (Victora et al. 2015) and is more likely the result of potentially several risk factor persistent for longer period or even throughout pregnancy, such as chronic malnutrition (Victora et al. 2015).

The role of immaturity at birth as a risk factor for later life ill health has been reviewed recently by for example Belbasis et al. (2016) and Raju et al. (2017). While the first one was focused on birth weight and the latter one on preterm birth, the studies identified in the review did not always distinguished between LBW and PTB. Immaturity was associated with adverse effects on several organ systems (Table 3). The reported effects range from structural abnormalities (e.g. reduced grey matter brain volume; abnormal growth/maturation/function of vascular tree), to increased prevalence of ICD-10 recognized diseases (e.g. asthma, type 2 diabetes, depression). Immature newborns have increased disease susceptibility throughout the life, even if they are born otherwise healthy.

<table>
<thead>
<tr>
<th>Organ system</th>
<th>Health effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central nervous system</td>
<td>Decreased intelligence; structural abnormalities; psychiatric diagnosis</td>
</tr>
<tr>
<td>Metabolic</td>
<td>Type 2 diabetes mellitus; metabolic syndrome; increased total cholesterol</td>
</tr>
<tr>
<td>Respiratory</td>
<td>Asthma; pneumonia; low respiratory infection reduced lung function</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>Elevated blood pressure coronary heart disease; cerebrovascular disease; abnormal growth/maturation/function of vascular tree</td>
</tr>
<tr>
<td>Other</td>
<td>Mortality; childhood stunting</td>
</tr>
</tbody>
</table>

**Preterm birth (PTB)**

Preterm birth is defined as gestational age of <37+0 weeks at birth (WHO 2004). It is further subdivided into extremely preterm (<28+0), very preterm (28+0 to 31+6) and moderate to late preterm birth (32+0 to 36+6). According to ICD-10 term birth is defined as a birth between 37+0 weeks to 41+6 weeks. The Society for Maternal-Fetal Medicine and the American College of Obstetricians and Gynaecologists further divide term birth into early (37+0 to 38+6), full (39+0 to 40+6), late term (41+0 to 41+6) and post term (42+0) (Quinn et al. 2016), to reflect the observation that early term newborns have higher risk for long-term complications than their full term peers (Chawanpaiboon et al. 2019). No lower limit for extreme PTB, which would be the cut-off between spontaneous abortion and viable birth, has been agreed internationally (Quinn et al. 2016). In Finland, the threshold is set administratively by the recording of a birth in the Medical Birth Register (MBR). Live and stillbirths of at least 500g or 22 weeks gestation are recorded (THL 2019).

Antepartum estimation of gestational age relies either on ultrasound measurements of fetal size or recall of the last menstrual period (Engle 2006). Estimation based on the date of the last menstrual period is usually inaccurate in a range of two weeks. Estimation based on ultrasound measurements of embryonic crown-rump length at 5 to
2 weeks gestation is more accurate with about 3 days uncertainty (Engle 2006). Ultrasound based estimates tend to be lower than those based on the date of the last menstrual period, thereby leading to slightly increased PTB assignment (Quinn et al. 2016).

PTB can occur spontaneous or be induced (Parets et al. 2014). Induced PTB are those deliveries by caesarean section or induced labour before 37 completed weeks of gestation for maternal or fetal indications (Parets et al. 2014), accounting for roughly a third of all PTB (Goldenberg et al. 2008). Spontaneous PTB can either occur as preterm labour with intact membranes, accounting for 40-45% of all PTB (Goldenberg et al. 2008), or as preterm rupture of membranes, irrespective of whether delivery is vaginal or caesarean section, accounting for 25-30% of all PTB (Goldenberg et al. 2008).

Preterm birth disrupts the normal organogenesis, especially in organs with branching morphogenesis, such as lung and kidney, leading to deficiencies in organ structure and function (Reichetzeder et al. 2016).

**Low birth weight (LBW)**

LBW defined as below 2,500 g is the most commonly used indicator for underdeveloped newborns, despite its low specificity in detecting pathologically small newborns. The cut-off at 2,500 g is arbitrary without justification for its selection. While LBW as such is problematic due to its dependency on gestational age, it is easily available even in low resource setting and supports comparability across studies and populations (Hughes et al. 2017). It is still debated whether LBW as such, independent of gestational age and poor fetal growth, is on the causal pathway for mortality and other adverse effects. Some argue that it may be nothing more than a convenient surrogate for other factors associated with mortality and other adverse outcomes (Hughes et al. 2017).

LBW itself is a very broad indicator for abnormal prenatal development thereby a wide variety of underlying causes can lead to it. LBW can be the result of PTB of a newborn of appropriate size for gestational age. In this case LBW is a secondary effect of the underlying etiology of PTB, not a result of growth restriction. On the other hand, intrauterine growth restriction (IUGR) can lead to a too small a size at birth, indicated for example by LBW. Importantly, not all growth-restricted newborns necessarily are lighter than the cut-off for LBW and may therefore be categorised as “healthy” in epidemiological studies with LBW as the main endpoint (Hughes et al. 2017).

**Immaturity as indicator of intrauterine growth restriction**

Immaturity is an important, yet unspecific, endpoint in assessment of general health status in newborns. Body size is routinely used as indicator of physical development. Since the shortcomings of traditional LBW definition were recognised, efforts have been taken to identify and define better indicators for immaturity. A majority is still based on the comparison of anthropometric measures at gestational age to a reference population. However, the wide distribution of body size (in terms of body weight, body
length, head circumference) poses challenges in discriminating pathologically small newborns from constitutionally small newborns (Hughes et al. 2017).

Despite the consensus that low birth weight per se is not a meaningful endpoint to identify pathologically small newborns, no agreement has been reached on how to define a more meaningful one. While for epidemiological studies it may be sufficient to identify growth restriction at birth, in the clinical setting the early detection and monitoring of growth restriction is crucial for providing the best care and prognosis for the unborn child (Gordijn et al. 2016). This partly results in the different terminology of small for gestational age (SGA), which is measured at birth, and IUGR, which is diagnosed during prenatal ultrasound scans (Gordijn et al. 2016). In settings with regular antenatal care and regular ultrasound scans, physiological parameters and the growth rate can be taken into account when evaluating fetal growth. In settings with poorer health care, less ultrasound measurements may be available and the growth rate cannot be monitored throughout pregnancy. In such setting size at birth must serve as an indicator for prenatal growth rate. In retrospective epidemiological studies data on ultrasound measurements are often unavailable and proxies for growth restriction must be used based on newborn biometrics. Clear diagnostic criteria for growth restriction are yet to be agreed on. The lack of clear diagnostic criteria is reflected in heterogeneous terminology often used synonymously (Gordijn et al. 2016; WHO 2004). The terms IUGR and fetal growth restriction are often used synonymously in the literature.

In terms of diagnostic criteria, the wider agreement is that birth weight alone is not a sufficient criterion to identify pathologically small babies (Gorgijn et al. 2019; Sharma, Farahbakhsh et al. 2016; Victora et al. 2015). For IUGR consensus was recently reached that abdominal circumference and pulsatility index of umbilical artery and uterine artery should also be taken into account during ultrasound scans (Gordijn et al. 2016).

While body proportions are used in the diagnosis of IUGR (Gordijn et al. 2016), additional anthropometric measures beyond birth weight are not routinely described in the assessment of immaturity at birth. However, the importance of body proportions at and after birth as an indicator for growth restriction is recognized (Victora et al. 2015). Notably, the biometric parameters can be differently affected by growths restriction, characterizing different phenotypes. The etiology and risk factors differ between phenotypes and further division into specific growth restriction types is warranted (Victor et al. 2015).

The great variability and flexibility in human development, body size and gestational age remains the main challenge in differentiating immature, pathologically small newborns from constitutionally small newborns. Human development is a multidimensional continuum and defining thresholds to define “abnormal” development based on few easily assessed phenotypes will naturally lead to misclassification. Few studies include a combination of changes in body size and proportions after prenatal exposure to assess growth restriction in specific body compartments, which can result in differentiation similar to symmetric or asymmetric IUGR. It could potentially provide input for mechanistic toxicology studies. Data are
missing to reliably identify subtypes of immaturity, which may be associated with a different set of long-term health consequences.

### 2.3 MATERNAL SMOKING

#### 2.3.1 Composition of tobacco smoke and smoking rates

Tobacco smoke is a widely used stimulant, the health effects of which are well documented. Maternal smoking is a prime example of a detrimental, common exposure, which theoretically is easily modified by stricter tobacco laws and smoking cessation support. It is an excellent example exposure to investigate the potential of the DOHaD paradigm to explain disease burden, due to the well-studied effects in humans and animals.

Tobacco smoke is the product of combustion of dried tobacco leaves and additives. The combustion products are estimated to include roughly 7,000 different chemical compounds (Table 4) (U.S. Surgeon General, 2010). Smoking rates in the Finnish population vary by gender and education background. While smoking rates in working aged men developed positively with decreasing rates of daily smokers and increasing rates of never smokers, in 2016 only 46% of men were never smokers in a population based study. On the other hand, among working aged women there was no clear trend in daily smoking range, fluctuating between 14% and 20% from 1978 to 2016. Moreover, the proportion of never smokers decreased from 70% during 1978 - 1980 to 57% in 2016 (Ruokolainen et al. 2019). The socioeconomic differences in smoking rates increased among the Finnish working age population between the late 1970s and 2016. Higher education is strongly associated with lower smoking rates and vice versa (Ruokolainen et al. 2019). On average, about 10% of pregnant women smoke in Europe, however rates vary greatly among countries. The rates range from 5% in Lithuania to 22% in France (Banderali et al. 2015). In contrast to the other Nordic countries, no significant decline in maternal smoking rates was observed in Finland between 2000 and 2010 (Ekblad et al. 2013).
The Tobacco Act (549/2016) was introduced in Finland in 1977 and made Finland an example in the fight against the public health impact of tobacco smoking, as there was barely any regulation of tobacco products in other countries (Heloma & Puska 2016). Ban on advertising, sales to individuals younger than 16 years and smoking in schools and public places were the main provisions of the legislation. In addition, health warnings were required to be printed on cigarette packs and 0.5% of tobacco revenue was to be used for work to reduce smoking. The act has been updated in 1994 further tightening the advertisement ban, increasing the legal age for tobacco purchases to 18 year and introducing a ban on smoking in work places. The amendment of the Tobacco Act in 2010 made Finland once again a forerunner in the fight against tobacco. Since
then, the Tobacco Act explicitly sets the aim of a smoke free Finland until 2040 (Ministry of Social Affairs and Health 2016). A smoking rate of <5% is to be reached by, amongst other strategies, stepwise limiting tobacco sales and banning the display of tobacco products (Heloma & Puska 2016).

### 2.3.2 Increasing evidence for health effects in adults and children

The high rates of smoking are attributable to the addictive properties of tobacco smoke, caused by nicotine. Evidence is emerging that nicotine is not the only addictive compound in tobacco smoke, indicated by stronger addictive properties of tobacco-derived nicotine than pure nicotine. However, there is limited evidence of the addictiveness of other minor tobacco alkaloids (US Surgeon General 2010). Smoking addiction trajectories are determined by a combination of psychosocial, biological and genetic factors (US Surgeon General 2010). The strong addictive properties of nicotine stem from its binding to the acetylcholine receptor in the brain, with the midbrain dopamine area being especially important during addiction initiation (Dani 2015). Addiction to tobacco smoke is a process with multiple steps, which are characterized by different genetic components. Sharp and Chen (2018) estimate that attempts to quit smoking are up to 30% controlled by genetic factors, while roughly 54% of smoking relapse is attributable to genetic factors.

<table>
<thead>
<tr>
<th>Category</th>
<th>Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer</td>
<td>Cancer of oropharynx, larynx, oesophagus, trachea, bronchus, lung, stomach, liver, pancreas, kidney, ureter, cervix, bladder, colorectal, bone marrow (acute myeloid leukaemia)</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>Stroke, aortic aneurysm, early abdominal aortic atherosclerosis in young adults, coronary heart disease, atherosclerotic peripheral vascular disease</td>
</tr>
<tr>
<td>Respiratory</td>
<td>Chronic obstructive pulmonary disease, asthma, increased susceptibility for tuberculosis, pneumonia</td>
</tr>
<tr>
<td>Reproductive</td>
<td>Reduced fertility in women, ectopic pregnancy, male sexual function – erectile dysfunction</td>
</tr>
<tr>
<td>Eye</td>
<td>Cataracts, age-related macular degeneration</td>
</tr>
<tr>
<td>Other associations</td>
<td>Periodontitis, diabetes, hip fractures, rheumatoid arthritis, reduced immune function</td>
</tr>
</tbody>
</table>

The US Surgeon General Report published in 1964 was the first comprehensive analyses on the devastating effects of tobacco smoke (US Surgeon General 2010). It is now established that smoking causes cancer, most prominently in the pulmonary system and upper digestive tract area, as well as cardiovascular and respiratory diseases, such as chronic obstructive pulmonary disease (Table 5). Maternal smoking is associated with pre- and postnatal mortality. There is no threshold or safe level of tobacco exposure and the health outcomes are directly correlated with the duration and exposure levels of smoking. In addition, modifications of cigarette design and tobacco composition aimed to reduce toxic emissions, such as the use of filters and ventilation holes or cigarettes with reduced tar and nicotine content, were not reflected in reduced
adverse health effects. The common underlying mechanisms of health effects include DNA damage, inflammation and oxidative stress (US Surgeon General 2010).

Prenatal exposure to maternal smoking causes adverse birth outcomes, such as LBW, PTB and fetal growth restriction (Banderali et al. 2015). Maternal smoking increases the risk for perinatal death. Exposure during organogenesis leads most prominently to altered brain structure and function, as well as reduced alveolarisation (Banderali et al. 2015). In recent years data on long-term later life health effects of in utero tobacco exposure are emerging. Effects on lung function, ranging from increased risk of wheezing and airway hyper-responsiveness to asthma and bronchitis are well documented (Banderali et al. 2015; Jaakkola et al 2006). Maternal smoking is associated with increased overweight and obesity rates and higher blood pressure (Banderali et al. 2015). The scientific evidence for life-long effects and generally increased susceptibility for chronic diseases in later life due to prenatal tobacco exposure is increasing. A generally increased susceptibility for chronic diseases in later life is emerging (Banderali et al. 2015). Recent studies also suggest an intergenerational association of grand-maternal smoking on grandchild asthma (Accordini et al. 2018; Bråbäck et al. 2018).

Tobacco smoke and nicotine exposure have been shown in animal studies and biomarker studies in humans to lead to the epigenetic changes, most commonly studied as DNA methylations. It has been suggested that these effects may differ depending on the individual genotype (Alvarado-Cruz et al. 2018). While studies on the effect of the complete tobacco smoke mixture itself are rare, effects on health of specific tobacco components are slowly emerging. Epigenetic effects of nicotine, polyaromatic hydrocarbons, lead, cadmium and arsenic have been reported (Alvarado-Cruz et al. 2018).

2.3.3 Smoking and the placenta

In the past, the placenta was thought to be an organ protecting the fetus from exposures, but this view had to be revised. Nowadays it is evident that most chemicals cross the placenta, some may accumulate in the placenta and/or be metabolised in the placental tissue (Barr et al. 2007; Myllynen & Vähäkangas 2013; Myllynen et al. 2009). Tobacco compounds, such as nicotine, its main metabolite cotinine and heavy metals (cadmium, lead, methylmercury) have been detected in cord blood and meconium (Barr et al. 2007). This is not surprising, as many tobacco components have low molecular weight and high lipid solubility, therefore readily cross the placenta (Griffiths et al. 2014).

The physiological function of the placenta is the exchange of nutrients and waste product between mother and fetus (Barr et al. 2007). To facilitate this exchange, several mechanisms to cross the placenta evolved: passive diffusion, facilitated diffusion, active transport and filtration. The mechanisms of transport of a chemical are dependent on its physicochemical properties, for example molecular weight and lipophilicity, and the potential to bind as a ligand to transporter molecules (Myöhänen & Vähäkangas 2012; Myllynen et al. 2007).
Nicotine and its main metabolite cotinine pass freely across the placenta leading to a relatively higher exposure of the fetus than the mother (Jauniaux & Burton 2007; Lambers & Clark 1996). Additionally, a carrier-mediated mechanism for nicotine has been proposed based on in vitro tests in JAR human choriocarcinoma cell line (Zdravkovic et al. 2005). In human placental perfusion studies, placenta retained nicotine, thereby potentially leading to a prolonged effect on fetus by slow, subsequent release of retained nicotine (Pastrakuljic et al. 1998).

Maternal smoking is of great concern for developmental origin of disease due to the high number of fetuses exposed to a mixture of known harmful chemicals, which can disrupt several organ systems in active smokers.
3 AIMS

The overall aim of this doctoral dissertation was to use the Developmental Origin of Health and Disease (DOHaD) paradigm as a framework to conduct health risk assessment to estimate disease burden associated with developmental exposures. As an example exposure, maternal smoking was chosen. Its long-term effect on child’s health was studied linking register data with burden of disease methods.

The specific objectives were to

1. Establish and document a birth cohort suitable for register-based long-term follow up for effects of environmental exposures in Finland.

2. Characterise the smoking behaviour in Finnish pregnant women to identify determinants, and spatial and temporal patterns of smoking, as well as to evaluate effectiveness of past smoking cessation policies.

3. Analyse the effects of smoking during pregnancy on body size and body proportions at birth to identify sensitive windows of exposure and propose mechanistic interpretations of the observed effects.

4. Demonstrate the concept of DOHaD attributable burden of disease in the child by estimating the life-long disease burden in the child associated with maternal smoking directly and mediated via intermediate risk factors in a chained risk model.

5. Update the epidemiological evidence for an association between maternal smoking and childhood cancer to prepare the register linkage for long term follow up of the MATEX cohort.
4 MATERIAL AND METHODS

4.1 STUDY DESIGN

A register-based approach was combined with burden of disease assessment, partly built on data from the literature and partly based on specific parameters obtained from the Finnish population. The core MATEX cohort was identified from the Finnish Medical Birth Register (MBR). The cohort was used for exposure assessment of smoking during pregnancy and an epidemiological study of the association between maternal smoking and body size and body proportions at birth. Subsequently, the long-term health impacts were quantified using a DOHaD-based chained risk model, which includes effects mediated via the initial birth outcomes as risk factors from the epidemiological study.

4.1.1 Establishment of the medical birth register-based cohort

The MATEX birth cohort has been identified from the MBR, including births between 1st January 1987 and 31st December 2016 (THL 2019). The cohort includes all births of newborns with at least 500g body weight or 22 completed weeks gestation in Finland. Information was recorded by nurses and midwives during antenatal care visits with a standardized form. Information was available about the mother, the newborn, the delivery and potentially diagnoses made during pregnancy and up to 7 days after birth.

Information was available for the mother’s smoking status, sociodemographics, municipality of residence, and reproductive history. Gestational age, birth weight and length, head circumference and perinatal mortality were recorded for the newborn. Place and method of delivery, as well as potential pain relief was included in the dataset. Diagnoses are recorded as ICD codes (ICD-9 from 1987 to 1995 and ICD-10 1996 to 2016). Risk factors and treatments were available as binary variables (yes/no). The data content and variable definition has been updated between 1990 and 2004. The data content of the MATEX cohort is described in detail in Paper I.

The data were stored in the secured THL rokostat linux cluster environment, with GitHub version control. The person identifiers are pseudonymised and the key file to translate study IDs to PINs remains with the register holder.

4.1.2 Exposure assessment

Maternal smoking in Finland was studied based on smoking data available in the MATEX cohort. Information on smoking were self-reported by the expectant women.
and recorded by nurses and midwives during antenatal care visits (Paper I; Paper II).
Throughout the work, maternal smoking was classified based on the following categories
1. non-smoker,
2. quitted smoking during 1st trimester
3. continued smoking after the 1st trimester
4. missing
Births between 1987 and 1990 were excluded due to a change in the definition of maternal smoking information from the 1st January 1991 onwards.
Trends were analysed for temporal pattern (annual level) and spatial differences (municipality level). As potential determinants of maternal smoking maternal age, parity, and socioeconomic status have been analysed. Pregnancy as an opportunie time for smoking cessation was studied by comparing smoking status between 1st and 2nd pregnancies in the MATEX cohort.

4.1.3 Exposure time dependent changes in prenatal growth attributable to maternal smoking
The MATEX cohort was used to conduct an epidemiological study of the effect of maternal smoking on birth outcomes, which are risk factors for later life disease. These risk factors were included in the chained risk characterisation conducted within this work.
The study population consisted of all singleton births between 1st January 1991 and 31st December 2016 available in the MATEX cohort. Newborns with missing information on exposure, outcomes and co-variables, as well as a recorded diagnosis of congenital anomaly (ICD-10 codes Q00-Q99), were excluded.
The exposures of interest were quitted smoking during 1st trimester and continued smoking after the 1st trimester with no smoking as reference. The outcomes of interest included four groups (Table 6, Paper III):
1. Preterm birth (PTB)
2. Low birth weight (LBW)
3. Small for gestational age (SGA)
4. Body proportions
Anthropometric measures included were body weight, body length and head circumference at birth (Paper III; Table 6).
Table 6. Variables in epidemiological analysis of maternal smoking and growth restriction (Paper III).

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Definition</th>
<th>Smoking</th>
<th>No. of observations</th>
<th>Adjustment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preterm birth</td>
<td>Gestational age &lt; 37+0 weeks</td>
<td>quitted</td>
<td>1 210 410</td>
<td>maternal age, sex, parity, SES</td>
</tr>
<tr>
<td></td>
<td></td>
<td>continued</td>
<td>1 286 667</td>
<td></td>
</tr>
<tr>
<td>Low birth weight</td>
<td>Weight &lt;2500 g</td>
<td>quitted</td>
<td>1 170 187</td>
<td>maternal age, sex, parity, gestational weeks, SES</td>
</tr>
<tr>
<td></td>
<td></td>
<td>continued</td>
<td>1 328 221</td>
<td></td>
</tr>
</tbody>
</table>

Small for gestational age

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Definition</th>
<th>Smoking</th>
<th>No. of observations</th>
<th>Adjustment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight</td>
<td>&lt;10th percentile of weight at corresponding gestational week</td>
<td>quitted</td>
<td>1 210 048</td>
<td>maternal age, sex, parity, SES</td>
</tr>
<tr>
<td></td>
<td></td>
<td>continued</td>
<td>1 327 783</td>
<td></td>
</tr>
<tr>
<td>Crown heel length</td>
<td>&lt;10th percentile of length at corresponding gestational week</td>
<td>quitted</td>
<td>1 210 048</td>
<td>maternal age, sex, parity, SES</td>
</tr>
<tr>
<td></td>
<td></td>
<td>continued</td>
<td>1 327 783</td>
<td></td>
</tr>
<tr>
<td>Head circumference</td>
<td>&lt;10th percentile of head circumference at corresponding gestational week</td>
<td>quitted</td>
<td>573 343</td>
<td>maternal age, sex, parity, SES</td>
</tr>
<tr>
<td></td>
<td></td>
<td>continued</td>
<td>601 407</td>
<td></td>
</tr>
</tbody>
</table>

Body proportions

| Ponderal index PI          | >90th percentile of weight:length ratio at corresponding gestational week | quitted | 1 088 451           | maternal age, sex, parity, SES, weight z-score |
|                           |                                                                 | continued | 1 196 479           |                                                 |
| Brain-to-body ratio BBR    | <10th percentile of head circumference-to-body weight ratio at corresponding gestational week | quitted | 518 704             | maternal age, sex, parity, SES, weight z-score |
|                           |                                                                 | continued | 541 549             |                                                 |
| Head-to-length ratio HLR   | >90th percentile of head circumference-to-body length ratio at corresponding gestational week | quitted | 521 420             | maternal age, sex, parity, SES                  |
|                           |                                                                 | continued | 547 597             |                                                 |

$^a$ quitted: quitted smoking during 1st trimester; continued: continued smoking after the 1st trimester

SGA was defined as the anthropometric measure below the 10th percentile of the Finnish standard reference population at corresponding gestational age, sex and parity (Sankilampi et al. 2013). As a reference 10-90th percentile was used. Additionally, low birth weight (LBW) was defined as weight below 2,500 g and the healthy reference was 2,500 - 4,500 g (Paper III; Table 6).

As indices for body proportions ponderal index (PI), brain-to-body ratio (BBR) and head-to-length ratio (HLR) were included (Paper III; Table 6). PI was calculated using birth weight in grams and body length in cm. Over 90th percentile of the study population was regarded as high PI. BBR was calculated based on head circumference in cm and birth weight in grams. The brain weight was estimated based on the head circumference according to the National Institute of Neurological and Communicative Disorders and Stroke’s Collaborative Perinatal Project (McLennan et al. 1983). Under
10\textsuperscript{th} percentile of the study population was categorised as low BBR. HLR was calculated using head circumference in cm and body length in cm. Over 90\textsuperscript{th} percentile of the study population was regarded as high HLR. For all three indices of body proportions the healthy range (below the cut-off for PI and HLR; above the cut-off for BBR) was used as reference. Percentiles of indices for body proportionality were calculated for each gestational week separately.

Covariables were defined similar for the exposure assessment and the register-based epidemiological study. In the exposure assessment covariables were studied as determinants of smoking during pregnancy. In the epidemiological study covariables were included as confounding factors (Paper II; Paper III).

Parity was categorised as nulliparous (no previous births) and multiparous (one or more previous birth) (Paper II; Paper III). Socioeconomic status (SES) was categorized as upper white collar (upper level employees with administrative, managerial, professional and related occupations), lower white collar (lower level employees with administrative and clerical occupations), blue collar (manual workers) and others (farmers, self-employed, students, pensioners, no information) and missing based on the Finnish national classification of occupations (Statistics Finland, 2001) (Paper II; Paper III).

4.1.4 Life-long disease burden in the child of maternal smoking

A DOHaD-based chained risk characterisation was conducted to estimate the long-term health burden in the child associated with maternal smoking, focusing on burden of disease of chronic diseases. The risk assessment covered the chain from developmental exposure, in this work maternal smoking, to intermediate risk factors apparent (based on the epidemiological analyses in Paper III and Chapters 4.1.2 and 5.1.3) and then subsequently to diseases throughout the life course (Figure 4).

The exposure assessment conducted within the MATEX study was used to quantify the exposed population. The results from the register-based epidemiological study were included as intermediate risk factors and their corresponding risk estimates used for PAF calculations in the chained risk characterisation. In addition to the risk factors identified in the epidemiological study, namely PTB and LBW, childhood overweight and obesity were included. Endpoints associated with maternal smoking or the intermediate risk factors were identified from the literature (Paper IV).
Systematic umbrella reviews were conducted to identify health endpoints in the child associated with (i) maternal smoking, (ii) PTB and LBW, and (iii) childhood overweight and obesity. Search strategies and other details are presented in Paper IV. PubMed was searched up to November 2018 for meta-analyses reporting statistically significant risk estimates for the four above listed associations. Only NCD were considered eligible for inclusion. Endpoints were excluded if no corresponding background burden of disease estimate was available from the Global Burden of Disease 2017 database (GBD 2018 Results tool). In case of overlapping meta-analyses the newer meta-analysis including a higher number of original studies was included (Paper IV).

A systematic literature review and meta-analysis was conducted to clarify the evidence for a possible association between maternal smoking and cancer in early life before the age of 20 years. Search strategies and other details are presented in Paper V. PubMed and Web of Science were searched up until 1st June 2015 for original studies examining the association between maternal smoking and the risk for any types of cancer before the age of 20 years. The reference lists of identified articles, reviews and meta-analyses, were checked for relevant studies. The search included any language in PubMed. First, abstracts were screened to exclude animal studies, in vitro studies, reviews and meta-analysis, and commentaries. Then the full-texts of the remaining articles were evaluated using pre-defined inclusion criteria (Paper V).

### 4.2 STATISTICAL ANALYSES

#### 4.2.1 Exposure analyses

Trends and determinants of maternal smoking were analysed for the years 1991 to 2015 (n = 1,435,009). Changes of smoking status in consecutive pregnancies were analysed in 368,930 women whose first and second pregnancies were available in the MATEX cohort. Smoking rates were calculated in reference to the total number of
pregnant women and differences between subgroups were calculated using chi-square test. Spatial trends were analysed on municipality and region level using 5-year averages (Paper II). The analyses were conducted in Microsoft Excel 2010 and R statistical software.

4.2.2 Epidemiological analyses

As an indicator of epidemiological study power, minimum detectable Risk Ratios (RR) have been estimated for three outcome incidence rates (10%; 5% and 1%) using epi.cohortsize function in epiR package. The assumed input values were an exposure rate of 10%, a study power of 90% and a confidence interval of 95%. The study power was computed separately for different potentially interesting sub-populations of the MATEX cohort.

The association between maternal smoking and preterm birth, small for gestational age and body proportions was quantified in Finnish singleton births (Table 6). The association of maternal smoking with growth restriction was analysed in a subset of the MATEX cohort. Multiple logistic regressions were performed to estimate Odds Ratios (ORs) with 95% confidence intervals (CI). The regression models were adjusted for potential confounders (Table 6). The data were analysed using R Statistical software (version 3.4.3). The regression was adjusted for maternal age (continuous in years), gestational age (continuous in weeks), parity (nulli- or multiparous), sex (male or female) and SES (Paper III).

4.2.3 Chained burden of disease model

Disease burden expressed as DALY was used to characterize the risks in the child attributable to maternal smoking in a cross-sectional model. BoD has been estimated with established methods using population attributable fraction (PAF) (Hänninen et al. 2014). Background disease burden in Finland in 2017 has been retrieved from GBD results tool (GBD 2018 Results tool). The chained risk model was calculated stepwise: first the direct disease burden was quantified. The direct burden included the burden directly associated with maternal smoking or any of the intermediate risk factors. In a second step, the contribution of maternal smoking to the disease burden associated with the intermediate risk factors was estimated (Paper IV).

The PAF for each endpoint (PAFEP) was derived from the fraction of exposed population in the whole target population and the associated relative risks, identified from the MATEX study and the literature (as described in Equation 4 in Chapter 2.1.2). A separate PAFEP was calculated for each association, such as intermediate risk factor – later life disease or maternal smoking – later life disease. The attributable disease burden (aBoDdirect) was calculated combining population attributable fraction (PAFEP) with background disease burden (BoDEP) for each endpoint (Equation 5).

\[ aBoD_{direct} = PAF_{EP} \times BoD_{EP} \]
The chained assessment of the indirect impact of maternal smoking mediated via the intermediate risk factors was calculated by estimating the contribution of maternal smoking to the intermediate risk factors (Equation 6). In case one endpoint was associated directly with maternal smoking and in addition with an intermediate risk factor, only the direct impact was considered and not the health impact mediated via the sequelae to avoid double counting. The indirect disease burden (aBoD\text{indirect}) of maternal smoking was calculated as product of contribution of maternal smoking to the intermediate risk factors (PAF\text{MS}) and the disease burden attributable to the intermediate risk factor (aBoD\text{direct}). The fraction of exposed population was identified from Paper II and the literature (Table 7). The analyses were conducted in Microsoft Excel 2010.

\[ aBoD_{\text{indirect}} = PAF_{\text{MS}} \times aBoD_{\text{direct}} \]  

(6)

Table 7. Risk factor definition and population attributable fraction (PAF) estimation for the association between maternal smoking and risk factors in this work.

<table>
<thead>
<tr>
<th>Maternal smoking</th>
<th>Preterm birth</th>
<th>Low birth weight</th>
<th>Childhood overweight</th>
<th>Childhood obesity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Definition</strong></td>
<td>Tobacco smoking pregnancy</td>
<td>Gestational age &lt;37+0 weeks</td>
<td>Birth weight &lt; 2500g</td>
<td>BMI ≥ 25 at 12-18 years</td>
</tr>
<tr>
<td>Exposure fraction</td>
<td>7% only 1st trimester</td>
<td>4%</td>
<td>3%</td>
<td>14%</td>
</tr>
<tr>
<td>RR\text{a}</td>
<td>1.38</td>
<td>2.22 (cont smoking) 1.11 (quit smoking)</td>
<td>1.37</td>
<td>1.55</td>
</tr>
<tr>
<td>PAF\text{MS}</td>
<td>0.03 (only 1st trimester smoking)</td>
<td>0.08 (smoking throughout pregnancy) 0.01 (only 1st trimester smoking)</td>
<td>0.03 (smoking throughout pregnancy)</td>
<td>0.04 (smoking throughout pregnancy)</td>
</tr>
</tbody>
</table>

\text{a OR odds ratios used as rough estimates of relative risks; PAF}_{\text{MS}} population attributable fraction of maternal smoking}

4.2.4 Meta-Analysis of childhood cancer studies

In the meta-analyses, OR was used as a measure of the association between maternal smoking and the risk of childhood cancer. Pooled estimates and the corresponding 95% CI were calculated using either a fixed effect model or a random effect model. For studies, which only reported risk estimates for categories of cigarettes smoked per day, an estimate for the nominal smoking status (smoking vs. non-smoking) was calculated. The Mantel-Haenszel Method was used to calculate the pooled estimate in the fixed
effect model using the above introduced weight and the logarithm of each individual risk estimate (Mantel & Haenszel 1959). In the random effect model the DerSimonian and Laird Method was used to calculate the pooled estimate (DerSimonian & Laird 1986). In contrast to the fixed effect model it takes into account the variance between studies, and thus random effects model was chosen as the more reliable model if there was evidence for heterogeneity. The Generic Inverse Variance Method was used to assign a weight to each study included in the pooled estimate (Cochrane Collaboration. 2011a). In case data were lacking to construct a 2X2 table, the 95% CI was used to calculate the standard error (SE) and to estimate the variance. The pooled estimates were then created using the inverse of the study variance as the weight. Cochran Q statistics (p value < 0.1 as the level for significance) and Higgins et al. (2003) I² statistics were used to assess heterogeneity (Cochrane Collaboration. 2011b). Fixed effects or random effects model, were chosen accordingly. Publication bias was assessed by funnel plots based on Egger’s regression (Egger 1997). Additional details are presented in Paper V. The analyses were conducted in Microsoft Excel 2010 and R statistical software.

4.3 ETHICS APPROVAL AND REGISTER DATA PERMIT

In accordance with the Finnish Medical Research Act (1999/488) the MATEX study including the birth cohort identified from the MBR has been evaluated and approved by the ethics committee of the Northern Ostrobothnia Hospital District (EETTMK 44/2016; issued 18th April 2016). The right to use of register data held by the National Institute for Health and Welfare was granted under the document number THL/838/6.02.00/2016 (issued 22nd June 2016). The study design is fully register-based. The individual-level data are coded and the key for the coding is stored by the register holder. The dataset is stored on a secured network drive, which can only be accessed by authorised persons listed in the ethics approval and approved by the register holder. Due to the full register-based design of the study, no informed consent was required from the study participants according to the Finnish Personal Data Act 1050/2018.
5 RESULTS

5.1 REGISTER-BASED APPROACH FOR IDENTIFICATION EARLY MARKERS OF EFFECT AND LIFE-LONG FOLLOW UP

5.1.1 Baseline MATEX cohort for life-long follow up

To facilitate life-long follow up via register linkage, a register-based birth cohort was established. The MATEX cohort was extracted from the Finnish Medical Birth Register (MBR) including all birth from 1st January 1987 onwards with annual updates. Currently the cohort includes births up until 31st December 2016 summing up to about 1.8 million births. Out of all births, about 3% were multiple births.

In singleton pregnancies, the mean age when giving birth was 29 years (standard deviation (SD) 5.3 years) and mean body mass index (BMI) was 24 kg/cm² (SD 4 kg/cm²). The mean maternal age was higher in multiple pregnancies. The majority of women (59%) were multiparous with 34% expecting their 2nd child and 26% expecting their 3rd or later child. Most women were married (64%). Most deliveries (77%) were vaginal; however 10% of births were urgent or emergency caesarean section. The infant mortality including stillbirths was about 0.8%. The baseline characteristics of the MATEX cohort are described in more detail in Paper I.

Figure 5. Possibility for register linkage for life-long follow up after developmental exposure (modified from Paper I).
Data availability varied during the study period. Information on maternal weight and height, as well as head circumference at birth was only available for year 2004 or later. Similarly, data on ultrasound screenings, diagnostics and treatments were only recorded since 2004. Diagnoses were recorded according to ICD-10 classification since 1996 and are recorded if they concern pregnancy, delivery or the newborn. The data content is described in detail in Paper I.

The potential of life-long follow up of the MATEX cohort is supported by the long tradition of health registers in Finland. The personal identification number (PIN) of the mother and child are recorded in the MBR and virtually all health or administrative registers, enabling a direct linkage of records of a specific individual. In Finland, registers are available for primary health care visits, hospital discharges, medication and cause of death. These general registers are supplemented with specialised registers, such as for cancer, congenital anomalies and induced abortions (Figure 5). Health registers can be supplemented with administrative registers, for example for address linkage, dwelling characteristics, and social support, as described in Paper I.

The big study population of the MATEX cohort supports nested case-control studies, sibling study design and 3 generation investigation. The study power ranges from very high to reasonable, depending on the sub-population (Figure 6). Additionally, the MATEX cohort can be annually updated to include the previous year’s births, adding roughly 50,000 births yearly.

In preparation of potential register linkage, the usability of the MATEX cohort for studying the associations of maternal smoking and childhood cancer were explored. A meta-analysis of maternal smoking and childhood cancer has been conducted to update risk estimates as input for study power calculations. The results are described in more detail in Chapter 5.2.2. Although the MATEX cohort is of considerable size, birth
anomalies and childhood cancer are such rare outcomes, that the statistical power in an epidemiological analysis would be limited (Paper I).

5.1.2 Exposure characteristics

Spatial patterns and temporal trends in smoking behaviour of pregnant women were analysed to evaluate effectiveness of past smoking cessation policies and determinants of smoking where studied to identify susceptible groups (Paper II).

In singleton pregnancies, the national mean smoking rate was about 15% between 1991 and 2016. After a temporary maximum smoking rate of 16.3% in 2012, smoking rates during early pregnancy fell continuously to 13.8% in 2016, the minimum during the study period. The fraction of pregnant women who quitted smoking during the 1st trimester continuously increased during the study period from 1% to 7%. Neither economic crises nor tightening of the tobacco legislation had direct impact on the smoking rates in the following years (Figure 7A; Paper II).

Maternal age, socioeconomic status as well as parity were determinants for smoking behaviour during pregnancy. Generally, smoking rates were higher in nulliparous, younger pregnant women, and pregnant women of lower socioeconomic status. Aggregated over the observation period, the highest smoking rate was observed in nulliparous, blue collar women, who were younger than 25 years old. Contrary, the lowest smoking rate was observed in multiparous, upper white collar women, who were older than 25 years (Figure 7B). Smoking rates were lower in multiparous women than nulliparous women, except of in the group of younger upper white collar. The fraction of women who quit smoking during the 1st trimester is almost double as high in nulliparous, than in multiparous women (Paper II).

Smoking rates during pregnancy differed on municipal and regional level. On average, smoking rates were highest in Päijät-Häme (19.6%) and lowest average smoking rates were found in Åland (Ahvenanmaa) (11.7%) (Figure 7C). Similarly, the temporal trends differed greatly by municipality. While smoking during early pregnancy decreased in Northern Finland, it increased in Eastern Finland. For example, smoking during early pregnancy decreased from 30% and 28% to 13% and 12% (5-year mean, 1991-1996 and 2011-2015) in Kittilä and Inari, respectively. In comparison, during the same time, maternal smoking rate increased from 17% to 35% in Outokumpu. On regional level, the increase was highest in the regions of Northern and Southern Savo, while the decrease was strongest in Uusimaa and Kanta-Häme. Smoking after the 1st trimester decreased in all regions between 1991 and 2016. Trends on municipality-level are described in Paper II.

In order to study the difference in smoking rates between nulli- and multiparous women, we compared smoking status during the 1st and 2nd pregnancy for all women with those pregnancies recorded in the MATEX cohort (n= 368,406). The clear majority (94%) of pregnant women, who did not smoke during their 1st pregnancy, reported to be also non-smoking during their 2nd pregnancy. Among those women, who smoked at
the beginning of their 1st pregnancy, 41% reported to be non-smoking throughout their 2nd pregnancy (Paper II).

Figure 7. Smoking during singleton pregnancy in Finland (1991-2016) (A) temporal trend with economic crises and legislative changes indicated; (B) by socioeconomic status, age group and parity; (C) spatial stratification on region level. (modified from Paper II).
5.1.3 Growth restriction attributable changes in body proportions at birth

Differences in sensitivity during time windows of prenatal development were studied in an epidemiological analysis comparing effects on preterm birth, body size and body proportions in newborns exposed only during the 1st trimester and those exposed also after the 1st trimester. Mechanistic interpretations for the observed differences can suggest the potential of persistence of the observed effects throughout the life course, as described in Chapter 6.1.2 and Paper III.

The unadjusted mean gestational age and anthropometric measures were similar in the newborns non-exposed and exposed to maternal smoking, except for birth weight (Paper III Supplement). Birth weight decreased with the duration of smoking with a mean 30 g less in newborns only exposed early and the mean weight almost 200 g lower in newborns exposed throughout their prenatal period.

Definition of endpoints to assess altered body size and body proportions that would allow the identification of sensitive exposure windows and mechanistic interpretations was challenged by the complex relationship of gestational age and the three available biometric measures (Figure 8). Body proportions change during prenatal development and therefore it was decided to study each biometric measure according to gestational age separately and three indicators for body proportions also by gestational age. The three included indexes of body proportion take into account the possible combinations of the three biometric measures. LBW, as defined by ICD-10 P07 code (<2,500 g), was identified as a poor indicator of growth restriction, since it identifies the great majority of preterm newborns, which have an appropriate weight for their gestational age according to reference growth charts, as growth restricted. At the same time, few late term newborns are categorised as immature based on LBW alone, indicating low specificity to distinguish preterm infants from those with decreased prenatal growths. The use of biometric measures by gestational age decreased the impact of gestational age at birth on the assignment of healthy or pathologically small. It shows that several anthropometric measures at birth are needed to evaluate healthy or restricted prenatal growth.
Smoking during early pregnancy did not affect preterm birth (<37 weeks). In contrast, the risk for preterm birth was clearly increased in women smoking after the 1st trimester (Figure 9; Paper III).

Any smoking during pregnancy was associated with an increased risk for LBW (<2,500g) (Figure 9). Smoking throughout pregnancy more than doubled the risk for LBW, while smoking only during early pregnancy increased the risk by about 10% in comparison to non-exposed newborns. The risks for low birth weight for gestational age (<10th percentile) were slightly lower, but comparable to the ORs observed for LBW defined as below 2,500 g (Paper III).

In order to study the underlying mechanism for the reduction in birth weight in more detail, the effect on birth length and head circumference at birth were analysed.
Body length (measured as crown heel length) was more susceptible to the effects of maternal smoking than body weight (Figure 9). The risk for a small head circumference for gestational age was elevated, although not statistically significant, for smoking during early pregnancy, and clearly increased in newborns exposed to smoking after the 1st trimester. Overall, smoking only during early pregnancy was associated with decreased body size (<10th percentile), albeit the effect size was considerably smaller than the effect size for smoking after the 1st trimester (Paper III).

Maternal smoking was associated with altered body proportions at birth expressed as a higher risk for high PI (>90th percentile), low BBR (<10th percentile) and high HLR (>90th percentile) (Figure 9; Paper III). Differences between the observed risks for altered body proportions associated with smoking during early and late pregnancy were much smaller than those for body size. Confidence intervals of risk estimates were overlapping in the case of high PI and low BBR. Risks for altered body proportions associated with smoking after the 1st trimester were lower than those for small body weight.

These results suggest that effects on body size and proportion at birth associated with exposure during early prenatal development are persistent and no catch-up growth to the physiological full potential of growth occurs. Additionally, it stresses the sensitivity of especially lean body mass during early development.

Figure 9. Effect of smoking during early and late pregnancy on gestational age, body size and proportions at birth (modified from Paper III). Black diamonds mark Odds ratios (ORs) of smoking quitted during first trimester and grey dots mark ORs of smoking continued after 1st trimester.

As there may be a relation to lifestyle in pregnancies of low maternal age and high parity or vice versa, a sensitivity analyses with an alternative adjustment model including a combined variable for maternal age and parity has been conducted. The resulting ORs of this alternative adjustment model did not differ from the reported ORs (data not shown).
The results were robust against stratification by socioeconomic status and birth year (Paper III Supplement). Additionally, inclusion of additional adjustment factors, such as maternal weight and height, co-morbidities, marital status and assisted pregnancy did not change the results (Paper III).

5.2 LIFE-LONG DISEASE BURDEN OF DEVELOPMENTAL ORIGIN

5.2.1 Health effects of maternal smoking and intermediate risk factors

As the epidemiological study showed, insults of maternal smoking during early pregnancy have persistent effect after the exposure ceases. In addition, LBW and PTB are associated with higher prevalence of diseases later in life. A chained risk model was developed to assess the burden of chronic diseases during the whole postnatal life associated with maternal smoking directly and mediated via adverse birth outcomes (LBW, PTB) and childhood obesity.

An umbrella review, during which of 3,766 articles were screened in three systematic PubMed searches, resulted in the identification of 27 eligible meta-analyses used for burden of disease assessment in a chained risk model. Most meta-analyses were identified for endpoints associated with PTB and LBW (n=8 for each). Overall, associations with cancer, cardiovascular disease, mental health diseases, asthma and diabetes have been included in the chained risk model (Paper IV).

In addition to the 27 meta-analyses eligible for the BoD assessment, a great range of additional endpoints have been identified, which could not be included in the BoD assessment due to lacking data on the background burden (Table 8). Of great concern are the endpoints, which are early markers of disease, such as impaired glucose metabolism and decreased lung function, as well as risk factors for disease, such as high cholesterol, elevated blood pressure and low bone mineral density. Especially markers of cognitive and social functioning, such as lowered IQ and lower attainment in school or higher school absence do not represent diseases per se, but they are important contributors to later life health and economic success.
Table 8. Exemplary outcomes associated with maternal smoking and associated intermediate risk factors identified in the literature review. Endpoints/associations included in the BoD assessment in black; empty cells present that no statistically significant association was identified (modified from Paper IV; Paper V).

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Developmental exposure</th>
<th>Intermediate risk factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Maternal smoking</td>
<td>Preterm birth</td>
</tr>
<tr>
<td><strong>Congenital anomalies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiovascular system</td>
<td>Lee et al. 2013; Nicoletti et al. 2014</td>
<td></td>
</tr>
<tr>
<td>Central nervous system</td>
<td>Hackshaw et al. 2011</td>
<td></td>
</tr>
<tr>
<td>Isolated cleft lip (with)out cleft palate</td>
<td>Hackshaw et al. 2011; Little et al. 2004; Xuan et al. 2016</td>
<td></td>
</tr>
<tr>
<td>Cryptorchidism</td>
<td>Zhang et al. 2015</td>
<td></td>
</tr>
<tr>
<td>Craniosynostosis</td>
<td>Hackshaw et al. 2011</td>
<td></td>
</tr>
<tr>
<td>Digestive system</td>
<td>Nicoletti et al. 2014; Hackshaw et al. 2011</td>
<td></td>
</tr>
<tr>
<td>Hernia</td>
<td>Hackshaw et al. 2011</td>
<td></td>
</tr>
<tr>
<td>Face and neck</td>
<td>Nicoletti et al. 2014</td>
<td></td>
</tr>
<tr>
<td>Musculoskeletal system</td>
<td>Nicoletti et al. 2014; Hackshaw et al. 2011</td>
<td></td>
</tr>
<tr>
<td><strong>Cancer</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lymphoma</td>
<td>Paper V</td>
<td></td>
</tr>
<tr>
<td>Nervous system cancer</td>
<td>Paper V</td>
<td></td>
</tr>
<tr>
<td>Testicular cancer</td>
<td></td>
<td>Cook et al. 2010</td>
</tr>
<tr>
<td>Wilms’ tumour</td>
<td></td>
<td>Chu et al. 2010</td>
</tr>
<tr>
<td>Other cancers</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cardiovascular disease</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increased diastolic blood pressure</td>
<td>Parkinson et al. 2013</td>
<td>Mu et al. 2012</td>
</tr>
<tr>
<td>Hypertension</td>
<td></td>
<td>Li &amp; Xi 2014</td>
</tr>
<tr>
<td>Stroke</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atherosclerosis</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Cont. Table 8

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Developmental exposure</th>
<th>Intermediate risk factor</th>
<th>Childhood overweight/overweight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive / Mental</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decreased intelligence</td>
<td>Maternal smoking</td>
<td>Preterm birth</td>
<td>Low birth weight</td>
</tr>
<tr>
<td>Huang, Zhu et al. 2016</td>
<td></td>
<td>Kormos et al. 2014</td>
<td>Yu et al. 2010</td>
</tr>
<tr>
<td>Special education needs</td>
<td></td>
<td>Twilhaar et al. 2018</td>
<td></td>
</tr>
<tr>
<td>Lower education qualification</td>
<td></td>
<td>Bilgin et al. 2018</td>
<td></td>
</tr>
<tr>
<td>Increased school absence</td>
<td></td>
<td></td>
<td>An et al. 2017</td>
</tr>
<tr>
<td>Criminal/Deviant behaviour</td>
<td></td>
<td>Pratt et al. 2006</td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td></td>
<td>Wojcik et al. 2013</td>
<td>Mannan et al. 2016; Quek et al. 2017</td>
</tr>
<tr>
<td>Autism</td>
<td></td>
<td>Wang et al. 2017</td>
<td>Gardener et al. 2011</td>
</tr>
<tr>
<td>Attention</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deficit/Hyperactivity Disorder</td>
<td>Twilhaar et al. 2018</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respiratory</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pneumonia in childhood</td>
<td></td>
<td>Jackson et al. 2013</td>
<td></td>
</tr>
<tr>
<td>RSV-related acute LRI in childhood</td>
<td></td>
<td>Shi et al. 2015</td>
<td></td>
</tr>
<tr>
<td>Wheezing</td>
<td>Silvestri et al. 2015; Vardavas et al. 2016</td>
<td>Sonnenschein-van der Voort et al. 2014</td>
<td>Mebrahtu et al. 2015b; Azizpour et al. 2018; Flaherman &amp; Rutherford 2006</td>
</tr>
<tr>
<td>Metabolic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increased BMI</td>
<td>Oken et al. 2008; Rayfield et al. 2017</td>
<td>Silveira et al. 2008</td>
<td></td>
</tr>
<tr>
<td>Metabolic syndrome</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metabolic syndrome</td>
<td></td>
<td></td>
<td>Kim et al. 2017</td>
</tr>
<tr>
<td>Increased insulin resistance</td>
<td></td>
<td></td>
<td>Friedemann et al. 2012</td>
</tr>
<tr>
<td>Increased fasting insulin</td>
<td></td>
<td></td>
<td>Friedemann et al. 2012</td>
</tr>
<tr>
<td>Increased cholesterol</td>
<td></td>
<td></td>
<td>Friedemann et al. 2012; Umer et al. 2017</td>
</tr>
<tr>
<td>Decreased bone mineral density</td>
<td></td>
<td></td>
<td>van Leeuwen et al. 2017</td>
</tr>
</tbody>
</table>
### Cont. Table 8

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Developmental exposure</th>
<th>Intermediate risk factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Maternal smoking</td>
<td>Preterm birth</td>
</tr>
<tr>
<td>Younger age at menarche</td>
<td>Yermachenko &amp; Dvornyk 2015</td>
<td></td>
</tr>
<tr>
<td>Increased cortisol secretion</td>
<td>Pearson et al. 2015</td>
<td></td>
</tr>
<tr>
<td>Invasive meningococcal disease</td>
<td>Murray, Britton et al. 2012</td>
<td></td>
</tr>
<tr>
<td>Chronic kidney disease</td>
<td>White et al. 2009</td>
<td></td>
</tr>
<tr>
<td>Periodontal disease</td>
<td>Martens et al. 2017</td>
<td></td>
</tr>
<tr>
<td>Dental caries</td>
<td>Hayden et al. 2013</td>
<td></td>
</tr>
<tr>
<td>Traumatic dental injuries</td>
<td>Correa-Faria &amp; Petti 2015</td>
<td></td>
</tr>
<tr>
<td>Decreased paediatric quality of life</td>
<td>Ul-Haq et al. 2013</td>
<td></td>
</tr>
<tr>
<td>Increased need for social benefits</td>
<td>Bilgin et al. 2018</td>
<td>Bilgin et al. 2018</td>
</tr>
</tbody>
</table>

### Mortality

|                          | Marufu et al. 2015; Pineles et al. 2016 |                          |                |                                |
| Neonatal death           | Pineles et al. 2016                     |                          |                |                                |
| Perinatal death          | Pineles et al. 2016                     |                          |                |                                |
| Sudden infant death syndrome | Zhang et al. 2013 |                          |                |                                |
| Adult all-cause mortality | Zhang et al. 2013                     |                          |                |                                |

All associations listed here were statistically significant at 95% confidence in the reviewed meta-analyses.

#### 5.2.2 Evidence for an association of maternal smoking with childhood cancer

In order to clarify the evidence for an association between smoking during pregnancy and cancer in the child’s early life (<20 years), a systematic literature review and meta-analyses was conducted. Separate meta-analyses were conducted for 8 cancer types: lymphoma, nervous system tumours, bone cancer, soft tissue cancer, renal cancer, hepatic cancer, germ cell tumours and leukaemia. From initially 880 screened abstracts, identified from PubMed, Web of Science and reference lists of relevant articles, 62 articles were included in the meta-analyses. Results are presented as pooled ORs based on random effects model. Maternal smoking was associated with lymphoma (OR 1.21; 95% CI 1.05-1.39) and nervous system cancer (OR 1.09; 95% CI 1.02-1.17). The association with lymphoma was driven by a clear association with non-Hodgkin lymphoma (OR 1.27; 95% CI 1.07-1.48). The subgroup analyses of nervous system tumours resulted in a stronger association of maternal smoking with neuroblastoma.
than with central nervous system tumours (OR 1.32; 95% CI 1.03-1.69 and OR 1.10; 95% CI 1.00-1.22, respectively). No association with other nervous system cancer subtypes was found. The meta-analysis of studies reporting risk estimates for smoking during pregnancy and childhood leukaemia included 21 studies. The analysis suggests no association between maternal smoking during pregnancy and leukaemia in early life. For germ cell tumours, hepatic tumours, renal tumours, soft tissue sarcoma, and bone cancer a limited number of eligible studies (n ≤5 for each cancer type) were identified leading to less than 1,000 cancer cases included in each meta-analyses. For these cancer types no conclusion of an associated could be drawn due to a lack of statistical power. There was no indication for publication bias. The results are presented in more detail, including forest plots, in Paper V.

5.2.3 Burden of disease in the child attributable to maternal smoking and intermediate risk factors

In Finland in 2017 about 1.4 million DALY were attributable to NCD across all ages. Maternal smoking and subsequent intermediate risk factors were associated with health endpoints in the child, which in total amount to 28% (382,388 DALY) of the total BoD. The biggest contributors were cardiovascular diseases, including coronary heart disease, hypertension and stroke, diabetes and certain cancers (acute leukaemia, lymphoma, nervous system cancer, testicular cancer, kidney cancer, oesophageal cancer, liver cancer, multiple myeloma, pancreatic cancer, thyroid cancer) (Figure 10).

Figure 10. Burden of disease associated with DOHaD (maternal smoking, low birth weight, preterm birth, childhood overweight). Surface areas of the spheres are proportional to the disease burden in DALY; Dark grey circles: BoD attributable to disease groups; white circles: BoD attributable to DOHaD in the corresponding disease group (data from Paper IV).

Childhood cancer, congenital anomalies and asthma were identified as long-term diseases directly associated with maternal smoking, leading to a direct disease burden
of 170 DALY in Finland in 2017. About 3% of PTB, 4% of LBW and 3% of childhood overweight were attributable to maternal smoking, resulting in an additional 1,040 DALY indirectly attributable to maternal smoking via these intermediate risk factors (Figure 11). Overweight is the largest risk factor with roughly 28,000 DALY, followed by LBW (roughly 4,000 DALY) and PTB (roughly 1,400 DALY). Cardiovascular and mental and cognitive diseases, as well as asthma, diabetes and cancer were identified to be associated with PTB, LBW or childhood overweight in the long-term. About 2% of the total NCD burden (34,000 DALY) was attributable to these developmental determinants of health (Paper IV).

Figure 11. Attributable disease burden (34,000 DALY) by endpoint: (A) in absolute DALY; (B) fraction of the endpoint background burden (modified from Paper IV).
6 DISCUSSION

The overall aim of this thesis was to use the Developmental Origin of Health and Disease (DOHaD) paradigm as a framework to conduct health risk assessment of maternal smoking as an example of a developmental exposure. The effect of maternal smoking on size and body proportions at birth and life-long disease burden were quantified. To achieve the aims, a register-based epidemiological study in Finland in 1991 to 2016 was combined with extensive literature reviews and burden of disease methodology.

Maternal smoking was chosen as an example of developmental exposure since it is well characterised. Albeit it is self-reported, the reporting bias and non-disclosure rates have been previously studied (Männistö et al. 2016). Results of epidemiological studies on maternal smoking have been published for the Finnish population, enabling a partial, initial validation of the results reported here. The fact that maternal smoking is a modifiable lifestyle choice is making it interesting for public health interventions.

Maternal smoking rates were generally higher in younger, primiparous pregnant women and those of lower socioeconomic status. While smoking rates during early pregnancy remained fairly stable since 1991, smoking in later pregnancy decreased during the same period. In the last 5 years of the study period (2012-2016) there was also indication for a decrease in smoking during early pregnancy (Paper II). Previous studies of maternal smoking trends extended only until the year 2011, thus not reporting this positive development (Männistö et al. 2016; Ekblad et al. 2013). In this study, maternal smoking was associated with overall reduced body size and altered body proportions. While the effect on body size was stronger for continued smoking, altered body proportions at birth showed almost similar risks for quitted and continued smoking. Quitting smoking during early pregnancy reduced risk for PTB to background (Paper III). While the observed risk for LBW and PTB were well in line with the previously published studies (Räisänen, Sankilampi, Gissler et al. 2014; Räisänen, Gissler, Saari et al. 2013; Lamminpää et al. 2013), the effect of maternal smoking on body proportions have not been studied in the Finnish population before. In the long-term, prenatal exposure to maternal smoking is associated with increased disease susceptibility for cardiovascular, respiratory, and mental disease, as well as diabetes and cancer. Early developmental markers of effect, such as LBW, PTB and childhood obesity, are associated with the long-term disease risk. The long-term disease burden of maternal smoking was 170 DALY, with an additional 1,000 DALY mediated via the considered developmental risk factors. The total long-term disease burden of early developmental risk factors was roughly 34,000 DALY (Paper V). No BoD estimates based on a chained risk model resulting from for developmental origin of disease have been reported previously.
6.1 REGISTER-BASED EPIDEMIOLOGY

6.1.1 Register linkage for life-long follow up

This work used a register-based approach, which is increasingly common in epidemiological studies, especially in DOHaD studies requiring long follow up periods. In this context a health register is defined as an “organised system with uniform data aimed at comprehensive coverage of target population with particular disease, condition or exposure” (Eloranta & Auvinen 2015). The introduction of a unique personal identification number in Finland in the 1960s supported the development of population based, individual level registers, which are linkable with each other (Thygesen et al. 2014). In Finland, registers are available for health and sociodemographic matters. Register-based cohorts provide opportune possibilities to study the developmental origin of health and disease (Lamminpää et al. 2017). Additionally, because of the long history of registers, a fairly long retrospective follow up is already possible. However, follow up of a birth cohort from the MBR, such as the MATEX cohort used in this work, is limited to a maximum of about 40 years currently, since the MBR was only established in 1987. This limits the possibility to study developmental exposures, early markers of disease and chronic disease diagnoses or death attributable to chronic disease, since the majority of chronic disease burden affects the older population. The register-based approach offered a big study population of almost 1.6 million mother-child pairs (Paper I; Figure 6), which provided sufficient study power to investigate changes in body size and proportions (Paper III). The MATEX cohort provides a good starting point for long-term follow up of the birth cohort as well as the investigation of other prenatal exposures, such as air pollution and chemicals (Paper I). The Finnish MBR has previously been used to study developmental origins of disease (Table 9). Several studies have used register linkages to study long-term effects of maternal smoking (e.g. Ekblad et al. 2017; Jaakkola & Gissler 2004; Leivonen et al. 2016).

Register data are collected routinely and not for a specific research question (Thygesen et al. 2014). As a consequence, confounder data is not always available and proxy variables have to be used instead. In this work, the lack of information on general lifestyle during pregnancy, including but not limited to alcohol consumption, exposure to second hand tobacco smoke, and use of nicotine products other than cigarettes, are of concern. Furthermore, there is limited control over the data collection overall, thereby data content may not be sufficiently detailed for specific research questions. Data collection itself is hard to follow and coding may differ between persons and institutions sending the data to the register. In the case of the Finnish MBR data content changed from the introduction in 1987 compared to the current data content (Paper I). While some variables were added to the MBR during the study period, such as head circumference and maternal height and weight, others were temporarily excluded, such as APGAR score at 5min. Thus, sub-population of the MATEX cohort limited by time
(2004-2016) had to be used for any analyses including data on head circumference (Paper III). This halved the available study population from 1.5 million singletons to 700,000 singleton births. Nevertheless the study population was sufficient for excellent study power (Chapter 5.1.1; Figure 6). Crucially, variable definitions may change requiring exclusion of some years due to incompatibility of data. In this work, the years 1987 to 1990 had to be excluded due to cigarettes per day as categories for maternal smoking and not divided into early pregnancy and late pregnancy. While the MBR contains some lifestyle related information, such as maternal smoking and body mass index, other health registers may hold even less information on lifestyle.

Table 9. Examples of studies utilising Finnish MBR data.

<table>
<thead>
<tr>
<th>Interest</th>
<th>Exemplary references</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(neuro) development and mental health in the child</strong></td>
<td>Ekblad et al. 2017; Girchenko, Lahti-Pulkkinen et al. 2018; Girchenko, Tuovinen et al. 2018; Malm et al. 2012; Polo-Kantola et al. 2014; Sucksdorff et al. 2018</td>
</tr>
<tr>
<td><strong>Prenatal exposures</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Maternal age</strong></td>
<td>Räisänen, Sankilampi et al. 2014; Räisänen, Kramer, Gissler, Saari, Hakulinen-Viitanen et al. 2014; Räisänen, Kramer, Gissler, Saari &amp; Heinonen 2014; Tran et al. 2013</td>
</tr>
<tr>
<td><strong>Maternal co-morbidities</strong></td>
<td>Girchenko et al. 2018; Heinonen et al. 2018; Koivinen et al. 2017; Leppävirta et al. 2017, 2018; Simoila et al. 2018; Turunen et al. 2019</td>
</tr>
<tr>
<td><strong>Medication use</strong></td>
<td>Malm et al. 2012; Rodriguez et al. 2019</td>
</tr>
<tr>
<td><strong>Sociodemographic indicators</strong></td>
<td>Bastola et al. 2019</td>
</tr>
</tbody>
</table>

In DOHaD research long-term follow up is required, but there is a gap in available information on exposure of the child after birth, for example to second hand smoke. Both second hand and active smoking, as well as other lifestyle factors are established risk factors for chronic disease development and therefore should be included as confounder in epidemiological analyses. In most health registers only ICD-10 diagnostic codes are recorded and not subtle changes that may indicate early markers of disease. Thereby studies on the disease development trajectories requiring early markers of
disease are complicated in fully register-based designs. A combination of register design and traditional epidemiological studies can at least partly circumvent this problem. The study population and basic information may be identified from the MBR and subsequently supplemented with questionnaire or measurement data, which have been collected specifically for the study. This approach was applied in Finland by for example Bell and colleagues (2019) and Jokela and co-workers (2018).

The question of data misclassification and handling of missing data has been discussed (Thygesen et al. 2014). Register-based research encourages discussion about the suitability of traditionally used statistical methods. The availability of vast datasets may be misused for data dredging and misleading post hoc analyses. Furthermore, in big datasets differences may result in statistically significant tests simply because the tests are not suitable for such big datasets (Thygesen et al. 2014).

Individual rights for privacy may conflict with the overall benefit of public health (Thurston et al. 1999; Miller et al. 2008). In Finland, register-based research does not require explicit informed consent from study participants. However, it does require positive review by an ethics review board and the register holder (Ludvigsson et al. 2015). Collecting explicit individual informed consent for the whole MATEX cohort would require immense resources, both time-wise and money-wise. This naturally does not free us, the researchers, from ethically good conduct. As explained in Chapter 4.1.1, the MATEX data is coded with study IDs to protect privacy of individuals. In Finland, the public attitude was found to be positive or neutral about register research in Finland. Eloranta & Auvinen (2015) report, that while the biggest concern was the privacy protection and personal information ending up in wrong hands, the majority of the general public views the benefit for public health as more important than the individual right to privacy.

6.1.2 Mechanistic interpretations

As described in Chapter 2.2.4, LBW as an endpoint in epidemiological studies is a general indicator for immaturity at birth, but cannot inform about whether the development is appropriate for gestational age or which type of growth restriction may contribute to the low birth weight. It was hypothesized that careful consideration of reduction in anthropometric measures other than weight and changes in body proportion can support similar interpretations as ultrasound scans for symmetric and asymmetric IUGR. In addition, it can support mechanistic interpretations of growth restriction.

Prenatal development has been characterized by three phases of cell growth: first a phase of cell division increasing the number of cells (hyperplasia), followed by a phase of hyperplasia and hypertrophy, the increase in cell volume, and then lastly a phase of predominantly hypertrophy (Walid et al. 2007; Sharma, Farahbakhsh et al. 2016). Prenatal development is a period of highly organized cell differentiation, migration and apoptosis. Any interference in these processes can lead to persistent changes in structure, function and susceptibility to diseases.
Newborns exposed to cigarette smoke only during the 1st trimester of pregnancy had slightly increased risk to be born light, short or with a small head circumference. Those, who were exposed also after the 1st trimester had double as high risk to be light or short and about 65% higher risk to be born with a small head in comparison to the non-exposed children (Paper III; Figure 9). This may seem reassuring for the long-term risks associated with smoking only during early pregnancy, possibly largely during a time when the women are not yet aware of their pregnancy. However, risk for altered body proportions, especially body length or head circumference in relation to body weight, are similarly affected when smoking during early or throughout pregnancy. This suggests potential of the fetus “catch up” on some of the growth, if maternal smoking is ceased early enough during pregnancy. However, the growth will remain behind that of non-exposed fetuses. It has been reported that maternal smoking during early pregnancy is not associated with changes in anthropometric measures in comparison to no smoking during pregnancy (Samper et al. 2012). In contradiction, this work demonstrates that smoking during early pregnancy is associated with reduction in weight, length and head circumference at birth (Figure 9). This stresses the sensitivity of early prenatal development to maternal smoking leading to persistent effects.

Body weight is the sum of lean body mass and fat body mass. Based on birth weight alone or it’s proportion to body length and head circumference it is difficult to interpret whether the reduction is mostly based on reduction of lean body mass, such as muscles and organs, or fat mass. It has been proposed that symmetrical IUGR, a result of early growth restriction, affects abdominal circumference proportionally to length and head circumference (Moh et al. 2012). As shown in this work, maternal smoking leads to asymmetrical growth restriction to the other end of the spectrum: more body mass for length, as indicated by an increased risk for a high PI. Additionally, a low BBR indicates that a decreased head size in relation to the body weight. Exposed newborns are at greater risk for smaller heads in relation to their body length, as indicated by the increased risk for a high HLR. Based on this, the reduction in body dimensions seems to be highest in head circumference and body length in relation to the overall body size. On the contrary, Samper and colleagues (2012) report a proportionate reduction in fetal body size after prenatal maternal smoking exposures. Still, they conclude that maternal smoking seems to have stronger effect on fetal lean body mass than fat body mass. This work suggests a similar effect. Body length and head circumference contribute towards the lean body mass. Thus a shorter body length and smaller head circumference lead to a reduction in lean body mass. It can be speculated that the observed high risk for low birth weight is at least partly attributable to a loss of lean body mass and secondly to loss of fat body mass.

The reduction in body size and changes in body proportions observed after exposure only during early pregnancy indicates an adverse effect on hyperplasia, leading to a decreased cell number and overall smaller fetus (Moh et al. 2012). If hyperplasia is affected, it seems likely that cell differentiation, migration and apoptosis may also be affected. The consequences range from major structural changes, such as congenital anomalies, to minor changes in function and susceptibility to diseases, which
may not be apparent until well into adulthood (Sharma, Farahbakhsh et al. 2016). The observed smaller head circumference correlates directly with a smaller brain volume. It has been proposed that nicotine exposure results in premature cell differentiation at the cost of cell division leading to cell death, structural changes and altered function (Moh et al. 2012; Ekblad et al. 2015).

Considering the complex processes during organogenesis, it can be expected that any insult has the potential to lead to persistent consequences. The evidence is clear that children born after IUGR, independent of exposure to maternal smoking, are more likely to suffer poor health throughout their lives (Sharma, Farahbakhsh et al. 2016). Early morbidity and mortality are both considerably higher, but even the survivors with no apparent consequences during childhood are at risk for poorer health later in life. Currently there is debate in the scientific community whether IUGR is a first symptom in the causal chain from development origin to later life disease or whether it is a cause itself of later life disease. Additionally, as discussed in Chapter 2.2.4, studies are lacking to be able to reliably distinguish healthy small newborns from pathologically small newborns and among the latter subgroups with different underlying causes, which may affect long-term prognosis.

6.1.3 Uncertainties of register-based analyses

The effects of maternal smoking of prenatal growth have been studied extensively, and are mostly in line with the results of this work. There is extensive evidence for a reduction of body weight in exposed newborns, both, based on fetal ultrasound scans (e.g. Abraham et al. 2017) and birth weight (e.g. Räisänen, Sankilampi et al. 2014; Inoue et al. 2017). Similarly there is accumulating evidence for shorter body length associated with maternal smoking exposure, either based on fetal femur length (e.g. Abraham et al. 2017) or body length at birth (e.g. Samper et al. 2012). A reduction in head circumference has been reported previously (Källén 2000; Samper et al. 2012; Abraham et al. 2017). Body proportions have been studied less extensively with contradictory results. This work reports an increased risk for higher PI at birth after exposure to maternal smoking. In contrast, Howe and colleagues (2012), report a statistically non-significant decrease in PI at birth in exposed newborns. No statistically significant effect of maternal smoking on PI was reported by Samper et al. (2012).

Maternal smoking has been studied using MBR data regularly. Some studies investigated determinants of maternal smoking (Härkönen et al. 2018; Räisänen, Kramer, Gissler, Saari, Hakulinen-Viitanen et al. 2014), while others investigated the health effects in the child (Lamminpää et al. 2013; Ekblad et al. 2017; Tran et al. 2013; Leivonen et al. 2016). Previous work by Räisänen, Sankilampi et al. (2014) was updated in this work by extending the birth cohort by 6 years, and inclusion of new endpoints indicating growth restriction. The extension of the birth cohort by roughly 240,000 births does not substantially change the study power for the well-established endpoints of LBW, PTB and SGA (weight <10th percentile) (Figure 6). However, it was possible to include head circumference, which was added to the MBR only in 2004, as an endpoint
with sufficient study power. Furthermore, our study was the first to report markers of body proportions in the Finnish population. The risk estimates for PTB (<37 weeks) reported in Paper III and Figure 9 are similar to those reported by Räisänen, Sankilampi et al. (2014). The risk estimates observed for LBW (<2,500 g) are higher in this work than those reported by Räisänen, Sankilampi et al. (2014), which may be explained by differences in the adjustments taken into account in the regression model. The risk estimates reported for low birth weight for gestational age are not directly comparable due to difference in the definition of SGA with a cut-off at 10th percentile in this work and at <2 SD used by Räisänen, Sankilampi et al. (2014). To date, no risk estimates for reduced body length or head circumference and altered body proportions have been reported for the Finnish population in addition to Paper III.

The smoking information contained in the MBR has previously been validated by serum cotinine samples collected during early pregnancy. In almost 10,000 samples collected between 1987 and 2011, the non-disclosure rate of smoking was 7%, and in 4% of women reporting smoking in the MBR, the cotinine serum concentrations were very low (Männistö et al. 2016). The MATEX cohort, analysed in this work, does not contain information on paternal smoking resulting in exposure of the mother to second hand tobacco smoke or information of smoking amount, such as cigarettes per day, or possible smoking cessation after the 1st trimester. Passive smoke exposure during pregnancy is associated with increased risk for PTB (Cui et al. 2016) and lead to an upwards bias in our risk estimates. In light of the absence of association between smoking only during 1st trimester and PTB it seems unlikely that SHS exposure confounded the results strongly. Paternal smoking has not been found to be statistically significantly associated with LBW in previous studies (Salmasi et al. 2010). It is possible that women who reported smoking after the 1st trimester ceased smoking later during pregnancy. This would be especially important exposure misclassification for the association with PTB, since it is an acute effect of smoking later in pregnancy. The observed spatial differences in smoking rates during pregnancy are not well understood, but may be partly due to differences in area level socioeconomic level and education (Räisänen, Kramer, Gissler, Saari, Hakulinen-Viitanen et al. 2014). Further studies are warranted to identify are level factors positively influencing smoking rates.

Tobacco smoke is a mixture of thousands of chemicals and so far there role of specific compounds is not sufficiently understood. Nicotine is rather well studied and mechanisms of the health effects have been identified. Nevertheless, nicotine replacement products are recommended by the Finnish Medical Society as a support for smoking cessation in women who are not able to quit smoking during pregnancy otherwise (Duodecim 2018). No information on nicotine exposure other than cigarette smoking was available, potentially leading to exposure misclassification, if nicotine is the tobacco compound responsible for many of the observed health effects.

No information on alcohol consumption during pregnancy was available. Prenatal exposure to alcohol causes the fetal alcohol spectrum disorder, which is in mild cases characterized at birth by small size and weight. Some of the observed effects on birth size may be attributable to alcohol exposure. In a population based online sample of 572
pregnant Finnish women, 14% reported alcohol consumption after awareness of pregnancy. The great majority (78%) of them reported 1-2 units of alcohol during the entire pregnancy, with only 1% reporting 1-2 units per week (Mårdby et al. 2017). The prevalence of prenatal alcohol exposure is so low, that it seems unlikely that it would be sufficient to strongly bias the observed effects.

Ideally, the anthropometric measurements recorded at birth would be supplemented with the fetal measurements conducted during ultrasound scans, in order to have a better understanding of the growth trajectories leading to the observed growth restriction. Furthermore, a follow up of growth would be warranted. Detailed body scans to identify the most sensitive tissues for growth restriction leading to the strong decrease in birth weight potentially can give hints to the mechanisms behind the growth restriction, which could clarify the questions raised in Chapter 6.1.2. Mechanistic toxicity is needed to explain the observed effects in detail. A better understanding can support the identification of pathologically small children from constitutionally small ones, as well as planning of interventions aiming to reduce the long-term consequences of immaturity at birth.

Besides wrongly self-reported information of the mother, other data errors are possible. The anthropometric measurements are influenced by a random error, which does not bias the observed association in any direction. Gestational age estimates are not very precise, as discussed in Chapter 2.2.4, but the error is similar across the whole study population and thereby does not bias the observed association in any direction.

6.2 BURDEN OF DISEASE OF DEVELOPMENTAL ORIGIN

6.2.1 Potential of DOHaD to explain disease burden

A couple of decades ago there was great hope that advances in genetics would provide insight into the share of the background risk provided by individual genetics, and the share of modifiable risk factors, such as lifestyle. The ultimate hope was that understanding the genetic and the lifestyle components in disease development could be used for targeted intervention to reduce disease burden. Since then, recent advances in developmental origin of diseases and epigenetic studies, have highlighted the complexity of disease etiology and how little is understood well enough to use it for disease prevention. NCD are often associated with multiple risk factors and symptoms, which themselves may be associated with several diseases. DOHaD has the potential to explain some part of the current disease burden, but the mechanisms are not well understood, yet (see Chapter 2.2). Mechanistic studies are required to understand the pathway from developmental exposure to later life disease in detail. Nevertheless, there is general agreement that the developmental period is crucial for health throughout life. The World Health Organization’s (WHO) First 1,000 Days life-course approach, which recommends that there should be a “focus on early childhood development, but given
the life-long associations between prenatal factors, health and socioeconomic status, life-course actions must start at the preconception and pregnancy stages and be sustained throughout childhood” (Jacob et al. 2017).

Table 10. Estimation of burden of disease attributable to the developmental origin according to GBD2017 (GBD 2018 Result tool).

<table>
<thead>
<tr>
<th>Disease</th>
<th>Risk factor</th>
<th>Low birth weight and short gestation [attributable DALY]</th>
<th>Low birth weight for gestation [attributable DALY]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neonatal preterm birth</td>
<td>7,027 (5,303-9,365)</td>
<td>7,027 (5,303-9,365)</td>
<td></td>
</tr>
<tr>
<td>Sudden infant death syndrome</td>
<td>45 (25-79)</td>
<td>16 (8-29)</td>
<td></td>
</tr>
<tr>
<td>Upper respiratory infections</td>
<td>0 (0-0)</td>
<td>0 (0-0)</td>
<td></td>
</tr>
<tr>
<td>Lower respiratory infections</td>
<td>23 (17-31)</td>
<td>11 (8-15)</td>
<td></td>
</tr>
<tr>
<td>Otitis media</td>
<td>0 (0-0)</td>
<td>0 (0-0)</td>
<td></td>
</tr>
<tr>
<td>Diarrheal diseases</td>
<td>9 (6-12)</td>
<td>4 (3-5)</td>
<td></td>
</tr>
<tr>
<td>Neonatal encephalopathy due to birth asphyxia and trauma</td>
<td>663 (488-877)</td>
<td>366 (268-493)</td>
<td></td>
</tr>
<tr>
<td>Neonatal sepsis and other neonatal infections</td>
<td>163 (31-239)</td>
<td>81 (15-121)</td>
<td></td>
</tr>
<tr>
<td>H influenza type B meningitis</td>
<td>3 (2-5)</td>
<td>1 (1-2)</td>
<td></td>
</tr>
<tr>
<td>Encephalitis</td>
<td>2 (1-3)</td>
<td>1 (0-1)</td>
<td></td>
</tr>
<tr>
<td>Pneumococcal meningitis</td>
<td>2 (1-3)</td>
<td>1 (0-1)</td>
<td></td>
</tr>
<tr>
<td>Meningococcal meningitis</td>
<td>1 (1-1)</td>
<td>0 (0-1)</td>
<td></td>
</tr>
<tr>
<td>Other meningitis</td>
<td>2 (1-3)</td>
<td>1 (1-1)</td>
<td></td>
</tr>
<tr>
<td>Haemolytic disease and other neonatal jaundice</td>
<td>11 (8-16)</td>
<td>6 (4-9)</td>
<td></td>
</tr>
<tr>
<td>Other neonatal disorders</td>
<td>812 (610-1065)</td>
<td>437 (323-583)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>8,762 (6,494-11,699)</td>
<td>7,952 (5,935-10,626)</td>
<td></td>
</tr>
</tbody>
</table>

BoD assessment of developmental origin of disease, as conducted in this work (Paper IV; Chapter 5.2.3) has not been published previously. While the WHO emphasizes the importance of early life for life-long health, the lack of robust evidence and risk estimate for the long-term effects of developmental exposures and early risk factors is reflected in the burden of disease estimates in the GBD Study (for details see Chapter 2.1.1). In their assessment, LBW and short gestation are included as risk factors for disease burden and PTB as a disease endpoint. Interestingly, they mostly include infectious diseases affecting children, which are not included in the chained risk model presented in this work (Table 10). According to the GBD 2017 study, neonatal PTB was associated with the great majority of attributable disease burden. In the GBD2017 assessment, no attributable burden is estimated for long-term chronic diseases later in life. In this work presented here (Paper IV; Chapter 5.2.3), the LBW and PTB attributable burden of NCD has been estimated to be roughly 5,200 DALY. According to the chained risk model applied in this work, roughly 3% of PTB disease burden and 8% of LBW disease burden are attributable to maternal smoking (Table 7). The GBD study reports roughly 7,000 DALY as PTB attributable disease burden (Table 10), of which roughly 210 DALY are attributable to maternal smoking, according to the chained risk model presented in Paper IV and Chapter 5.2.3 of this work. Similarly, for LBW 640 to 700 DALY could be attributed to maternal smoking. Thus, in Chapter 5.2.3 reported estimation of the disease burden with developmental origin is likely to be an underestimation.
Maternal smoking can be considered to be a form of second hand tobacco smoke. The GBD study reports 71 DALY in <20 year olds and 13 DALY in <1 year olds attributable to second hand tobacco exposure in Finland in 2017 (GBD 2018 Result tool). In light of the DOHaD-based chained risk model this work (Paper IV), the GBD estimates of tobacco attributable disease burden in children and young adults underestimates the harm and disease burden attributable to tobacco smoke exposure in early life.

The potential effect of prenatal exposures on disease susceptibility or predisposition is generally agreed upon. Obesogens, environmental chemicals increasing the risk for obesity later in life, and endocrine disruptive chemicals have been linked to later life dysfunction. Low nutrition and the “Hygiene hypothesis”, highlighting the importance of exposure to foreign antigens for proper immune system development, are the other side of the same coin: exposures and conditions during developmental periods and their effects on disease predisposition later in life (Barouki et al. 2012). Research on developmental origin of disease has been discussed to be crucial in the identification of risk factors for disease development and thereby necessary for early prevention and intervention (Grandjean et al. 2015). Some even go as far as to propose that traditionally adult-associated chronic diseases, such as cardiovascular disease and diabetes, should be considered to be paediatric diseases as their origin lies in early life (Heindel & Vandenberg 2015). It does not seem unlikely that maternal smoking can increase the risk for disease later in life, leading to a considerable disease burden. Overall, more clinical and epidemiological studies are needed to understand the associations of prenatal exposures and disease trajectories in more detail, as well as toxicological studies to infer possible mechanisms of the effects.

A better understanding of the exposure-outcome relationship, such as by gender, age groups and subpopulations, as well as life table modelling are required to gain detailed understanding of the disease burden across the life-course attributable to maternal smoking and other developmental exposures. It is important to remember that BoD, as a concept, can only explain disease burden statistically at population level and cannot conclude whether a risk factor caused or contributed to disease onset in a specific individual.

Maternal smoking is clearly more prevalent in the population with a lower SES (Paper II) and thereby associated disease burden is likely higher in this population group than in those with a higher SES. This social inequality is potentially further deepened not only by individual SES but also neighbourhood level SES, as it was shown that maternal smoking was influenced by municipality level education, unemployment and income (Räisänen, Kramer, Gissler, Saari, Hakulinen-Viitanen et al. 2014). Additionally, it was shown that lower SES itself, independent of maternal smoking, is associated with higher prevalence of adverse birth outcomes (Räisänen, Gissler, Sankilampi et al. 2013; Räisänen, Kramer, Gissler, Saari, Heinonen 2014), while the risk estimates for reduced body size at birth were not sensitive to stratification by SES (Paper III). These socioeconomic differences in adverse birth outcomes are evident
6.2.2 Health endpoints in the causal chain

Maternal smoking was used as an example for a developmental exposure in this work to investigate the potential of DOHaD to explain disease burden. Concerns for the health effects of maternal smoking were first raised by the US Surgeon General in the 1960s, and since then it has been studied extensively (US Surgeon General 2010; US Surgeon General 2014). Surprisingly, many knowledge gaps remain and especially the understanding of long-term effects and their underlying mechanisms remains limited. This limited the inclusion of endpoints into the BoD assessment, probably leading to an underestimation of the attributable burden. The background burden, retrieved from the GBD Study (GBD 2018 Results tool), is a cross-sectional assessment of BoD in a given year and population. This is by far not optimal for DOHaD attributable disease burden quantification, since decades lie between the exposure and outcome.

The burden of disease quantification in this work assumes causality based on statistical significance in meta-analyses of observational studies (Paper IV; Chapter 4.1.4). This is of course a very simplified approach since statistical significance alone does not indicate plausibility and the direction of association. However, for some health endpoints early markers of disease have been reported (Table 10), for example hypertension and elevated blood pressure during childhood (Parkinson et al. 2013; Li et al. 2014; Umer et al. 2017; Friedemann et al. 2012; Llewellyn et al. 2016), and diabetes and increased insulin resistance and fasting insulin levels in blood (Llewellyn et al. 2016; Verbeeten et al. 2011; Friedemann et al. 2012).

The choice of developmental exposure and outcome under investigation is at least partly dictated by statistical power, thereby limiting studies to more common exposures and outcomes, such as indicators of immaturity, e.g. LBW and PTB (3% and 4% incidence rate respectively in the MATEX cohort), and cardiovascular disease and diabetes. Rarer endpoints, such as childhood cancer, are less well studied and evidence is limited despite understanding of some underlying mechanisms (Paper V). The use of immaturity as proxy for developmental exposure or sub-optimal prenatal conditions in epidemiological studies is problematic, as described in Chapter 2.2.4. Immaturity at birth is a symptom or outcome of a wide variety of complications during pregnancy, from poor nutrition to maternal smoking and infections. Depending on the cause of immaturity at birth the long-term outcomes may differ. With advances in register-based studies and big birth cohorts, it is necessary to stratify epidemiological analyses by underlying cause and compare long-term health outcomes and early markers of disease. Additional anthropometric indices, such as the body proportions proposed in this work (Chapter 5.1.3), can guide the identification of different sub-groups of immature infants with different associated long-term risks. This could shed light on underlying disease mechanisms. In the chained risk model applied in this work, these complications are not taken into account. Due to limited literature availability, a fraction of long-term
disease burden attributable to intermediate risk factors was assigned to the effect of maternal smoking, although there may have not been evidence for a direct association from maternal smoking via intermediate risk factor to health outcome.

6.2.3 Uncertainties in the chained risk model

The uncertainties in the chained risk model developed here (Paper IV; Chapter 4.2.3) arise from three levels: (i) the BoD methodology itself, (ii) the background disease burden retrieved from the GBD study and used as estimation basis in this work, and (iii) the epidemiologic evidence identified in the literature review. Due to the uncertainties in the model, the reported attributable disease burden must be understood as a rough, order of magnitude estimate and not as a numerically perfectly true estimate.

The BoD methodology is applied in a cross-sectional manner. It cannot accommodate the longitudinal dimension of developmental origins of disease. A long time of latency between the exposure and the outcome would warrant a life-table approach with modelling the disease burden across age groups and years. This was not possible due to the limited period for which the GBD background burden is available, namely 1990 to 2017. Thus, the exposure prevalence, intermediate risk factor and outcome disease burden were applied in a cross-sectional manner. This ignores changes in the population structure, such as increasing elderly population, and changes in the prevalence of NCD. In the chained risk model it was assumed that the elderly population in 2017 was exposed to the same rates of maternal smoking and intermediate risk factors as the newborns in 2017. Similarly, it is assumed that the birth cohort of 2017 will have the same disease prevalence and population structure in 60 years from now as the currently 60 year old population. This is clearly an oversimplification of the true situation. Still, the order of magnitude of attributable disease burden estimated in the chained risk model should still be valid. A further uncertainty in the BoD methodology is the estimation of the PAF in a chained risk model. As described in Chapter 2.1.2 there was not suitable estimation method proposed in the literature for a chained risk model aiming to estimate indirect effects. Therefore a new estimation method was proposed, which was not based on three partial PAFs, but on the estimation of a compound PAF based solely on the exposure-intermediate risk factor and intermediate risk factor-outcome associations. The proposed method is conservative in the fraction of intermediate risk factor burden it attributed to the developmental exposure. Therefore it seems unlikely that it would overestimate the attributable disease burden.

The BoD estimations in this work (Paper IV; Chapter 5.2.3) are fully based on background disease burden from the GBD 2017 study (GBD 2018 Results tool). These estimates are based on established methodology and consolidated efforts for high quality input data. Although the data for Finland have not been validated independently, the background burden data are generally considered reliable. Associations with specific health endpoints and in specific age groups have been
identified in the literature review. The literature-based health endpoints and age groups did not always totally overlap with the GBD disease definition or available age groups. However, rough age groups, such as children, teenager and adults have been considered in the BoD estimation (Paper IV). Additionally, only those endpoints from the literature have been included, for which an equivalent was available in the GBD background burden database. Very specific health endpoints, which were not available as background burden on that level of specificity (for example specific congenital heart malformations), have been excluded to avoid overestimation of disease burden (Table 8). A bottom-up approach including the use of independently calculated background disease burden and risk estimates is crucial to validate the here estimated disease burden attributable to developmental origins.

The backbone of the here demonstrated chained risk assessment (Paper IV; Chapter 4.2.3) was the identification of long-term health effects of developmental exposures in systematic literature reviews. Any systematic literature review, such as in Paper IV and Paper V, is sensitive, amongst others, to the quality of original papers, reporting bias and poor search strategies. The literature reviews for the chained risk model were limited to meta-analyses which are sensitive to their own set of biases and problems, especially due to the long time of latency between the exposure and the outcome (Drucker et al. 2016). There is some risk that not all identified and included associations are causal or that the magnitude of the effect changes with increasing evidence. Nevertheless, as a scoping effort to demonstrate the concept of DOHaD in BoD assessment the included associations are deemed valid. For analyses of more specific aspects of the DOHaD attributable disease burden a literature review and meta-analyses of original studies is highly recommended to target the resulting pooled risk estimates to age groups and endpoints for which a corresponding background burden is directly available.

6.3 CONCEPTUAL UNCERTAINTIES

This work assumes the validity of the DOHaD paradigm. The underlying idea is that exposures during prenatal development interfere with healthy growth and that adverse birth outcomes are the first symptoms of impaired disease resistance throughout life. For many associations a causal pathway from developmental exposure via intermediate risk factors to long-term health effects is assumed although only partial evidence is available epigenetics and other mechanisms potentially mediating the effects.

An equally valid question in the DOHaD paradigm is the role of paternal factors before conception. The main research focus is on maternal factors, which is understandable because the fetus is directly depending on her, but studies start to accumulate that support the notion of the importance of paternal factors during the preconception period (Soubry 2018). This part of the equation remained fully unexplored in this work.
It is worth considering how the DOHaD paradigm may affect pregnant women. If most ill-health in the end boils down to sub-optimal conditions during developmental periods on top of genetics, it puts immense pressure on the expecting mothers to provide optimal developmental conditions, possibly even leading to feelings of guilt for any ill-health in the child. Even with harmful and avoidable exposures such as maternal smoking, putting guilt for “irresponsible behaviour” on women will not solve the problem, but only lead to hiding of socially un-accepted behaviour. In contrast, early education in health literacy can establish healthy lifestyles before pregnancy and risks can be avoided and do not need to be mitigated. It is the responsibility of the whole society, health care system and education system to raise awareness of risks associated with unhealthy lifestyle, not only for ourselves, but for the generations to come. With this responsibility comes the task to focus research on the disease prevention and not only treatment once the first symptoms occur.

Maternal smoking was chosen as an example for a well characterised prenatal exposure. The relevance of maternal smoking as an important risk factor in the future may be debated in the light of the legislative efforts and the Savuton Suomi (Smoke free Finland) initiative to see Finland smoke free until latest 2040 (Ministry of Social Affairs and Health 2016; Suomen ASH ry 2019). It may well be that the exposure to tobacco drastically decrease until 2040, but as this work demonstrates (Paper III, Paper IV), the health effects will remain for decades to come in both prenatally exposed individuals as well as active and passive smokers. Therefore, tobacco smoke will remain as an important risk factor for disease even after exposure drops practically to zero. On a wider conceptual level, maternal smoking was only chosen as an example. Alcohol consumption, nutrition and other environmental chemicals can be expected to have comparable long-term effects and should be studied carefully (Heindel et al. 2017; Heindel & Vandenberg 2016; Schug et al. 2012; Lunde et al. 2016).

6.4 FUTURE PERSPECTIVES

WHO’s efforts in the First 1, 000 Days approach is a leading public health effort trying to utilise the DOHaD paradigm for health intervention. It should be included not only in the training of health care experts, but also in health literacy education in school (Jacob et al. 2017). Its focus on disease prevention throughout life can provide an opportunity for public health to step away from treatment and intervention when first symptoms occur or risk factors accumulate. If the DOHaD paradigm is endorsed in future public health initiatives, healthy lifestyle and prenatal care that starts at pre-conception, will be of increasing significance. Furthermore, the focus would shift from each individual’s behaviour to family and social environment due to its direct influences on the next generation’s health.

The DOHaD paradigm is receiving increasing attention and new –omics technologies and advances in epigenetics in addition to the more traditional approaches
and concepts are contributing to understanding the underlying mechanisms of effect. A truly multidisciplinary effort is needed to understand the developmental origins of ill-health and to make the new understanding useful in disease prevention and treatment. An exploding amount of data, methods and computing power provides great opportunities to change our understanding of health and disease and new challenges in putting all pieces of evidence into context to understand the full picture and identify data needs (Dennis et al. 2016). In future, elaborate combination of register-based epidemiological studies to identify promising long-term health outcomes, as elaborated in Chapter 6.1.1, should be combined with molecular epidemiological designs, in which biological samples for biomarkers and early markers of disease development can be analysed, in order to study causal pathways in disease etiology. Epidemiological studies must be supplemented with in vivo and in vitro studies to identify mechanisms of effect. Understanding of etiology and long-term prognosis of phenotype at birth, including birth weight and length, body composition and proportions, are crucial to identify pregnancies and newborns at risk of poorer long-term prognosis and to understand the pathway from developmental origin and early changes in physiology, function and structure that may lead to clinical disease onset in the long-term.

Data content of health registers should be regularly updated if new data needs arise. The Finnish MBR should be extended to include other developmental toxicants, such as nicotine replacement products and e-cigarettes, as well as second hand smoking and alcohol consumption. A great limitation of the MBR, and in this work, is the lack of information on nicotine exposure other than cigarettes, such as nicotine replacement products, e-cigarettes and snus. These and SHS exposure should be recorded in the MBR, in the best case including information on the daily amount used and possible gestational week of use cessation. Assuming that at least some of the observed effects on prenatal growth are attributable to nicotine, the official endorsement of nicotine replacement products during pregnancy by the Finnish Medical Association may suggest false safety in pregnant women. There is limited evidence for safety of nicotine replacement products, and nicotine in general, during pregnancy. The risks associated with it must be communicated and clearly explained to pregnant women. Further studies of nicotine exposure via nicotine replacement products and e-cigarettes are urgently needed. Additionally, the inclusion of information on the level of exposure, for example as cigarettes or unit of alcohol per day in the MBR is recommended. Optimally, the MBR would not only record maternal exposure, but also paternal information. These data are crucial to support the use of the MBR in research and to identify risk pregnancies during antenatal care visits.

Maternal smoking is a complex public health problem, which needs to be tackled with a variety of strategies. Smoking initiation must be avoided and studies are needed to identify factors that contribute to smoking uptake in different Finnish regions and age groups. Support for smoking cessation should be increased already before the first contact within the antenatal care. As demonstrated, even smoking during early pregnancy can cause persistent effects in the newborn, which stressed the importance of an already smoke free pre-conception period. The Savuton Suomi 2030 initiative is a
promising start to a smoke free future, but uncertainties remain on how the goal of no smoking in Finland within a decade will be accomplished (Paper II).
7 CONCLUSIONS

Maternal smoking is associated with increased risk for adverse birth outcomes and life-long health consequences, contributing to some extent to the rise in chronic disease burden. The sensitivity of early prenatal development to tobacco smoke stresses the importance for smoking cessation before conception to avoid persistent health effects. This work demonstrates the importance of developmental periods for life-long health, which should be considered in a comprehensive risk assessment.

The long tradition in Finnish health registers provides the needed means to design register-based epidemiological studies with sufficient follow-up time and statistical power to study prenatal environmental exposures and associated health effects later in life. The functionality of the MATEX birth cohort, established from the Medical Birth Register, has been demonstrated and the usability for studying effects of prenatal exposures on pregnancy outcomes has been outlined.

The fraction of women quitting smoking during the first trimester rose steadily since 1991, indicating a better understanding of the risks associated with maternal smoking in pregnant women and diminishing social acceptance. Continued effort is needed to reach the goal of a smoke free Finland until 2030. Spatial differences in maternal smoking are not well understood and warrant further investigation.

The sensitivity of early prenatal development to tobacco smoke was associated with changes in body proportions, due to reduction especially in body length and brain development. The limited potential to repair damages induced in early pregnancy, marked by similar effects for smoking only during early pregnancy or throughout pregnancy, stresses the importance of a healthy lifestyle and avoidance of harmful exposure already before conception.

Later life consequences in the child are either directly associated with maternal smoking or mediated via intermediate risk factors. A wide range of later life effects, including respiratory, neurodevelopmental, mental, metabolic and cardiovascular diseases, are associated with maternal smoking and subsequent intermediate risk factors. Maternal smoking in Finland is associated with a loss of up to 1,200 years of healthy life. The chained risk assessment demonstrates the feasibility of DOHaD-based burden of disease calculations and sheds light into the largely unknown risk factors for disease susceptibility.

Overall, there is evidence that a portion of lymphomas and nervous system cancers could be attributed to maternal smoking. More studies investigating especially rare childhood cancers are needed. Better understanding of the possible mechanisms of childhood cancer development is needed to design targeted epidemiological studies.

Based on the example of maternal smoking this study demonstrates the sensitivity of fetal development to insults and their lasting effects, which may become obvious only later in life through increased susceptibility for disease. It demonstrates usability of the DOHaD paradigm in risk assessment, emphasizing the importance of healthy lifestyle and safe environment already before conception for a life-long healthy life.
Developmental origin of disease can and should be taken into account in health risk and impact assessment in order to reliably evaluate all associated loss of health. A register-based approach was shown to be a promising study design in life-course epidemiology, especially via register linkages and prolonged follow-up times.
8 REFERENCES


Jacob, C. M., Baird, J., Barker, M., Cooper, C., & Hanson, M. 2017. The importance of a life course approach to health: Chronic disease risk from preconception through adolescence and adulthood. *World Health Organization*.


Räisänen, S., Kramer, M. R., Gissler, M., Saari, J., Hakulininen-Viitanen, T., & Heinonen, S. 2014. Smoking during pregnancy was up to 70% more common in the most deprived municipalities - a multilevel analysis of all singleton births during 2005-2010 in Finland. *Preventive medicine*, 67, 6-11.


Sharp, B. M., & Chen, H. 2018. Neurogenetic determinants and mechanisms of addiction to nicotine and smoked tobacco. The European journal of neuroscience,


Yermachenko, A., & Dvornyk, V. 2015. A meta-analysis provides evidence that prenatal smoking exposure decreases age at menarche. Reproductive toxicology, 58, 222-228.


Chronic diseases are increasingly important in the developed world. It has been proposed that adverse exposures during pregnancy can harm the fetus and subsequently increase the susceptibility to develop chronic diseases later in life. This thesis aimed at using the Developmental Origin of Health and Disease paradigm as a framework to conduct environmental health risk assessment to estimate disease burden associated with maternal smoking as an example of a developmental exposure. The effect of maternal smoking during early and late pregnancy on body size and proportions at birth were investigated. Potential loss of healthy life in the Finnish population attributable to prenatal exposure to tobacco smoke was quantified.