Education and the basis of self-rated health

Liina Marjukka Junna
University of Helsinki
Faculty of Social Sciences
Sociology
Master's thesis
2016, February

Contents

1. Introduction 1

2. Measuring health 3

2.1 A working definition of health 3

2.2 How to measure health 6

2.3 Education as a determinant of health 10

3. Self-rated health (SRH) 11

3.1 Definition and recommendations 11

3.2 Psychometry 14

3.3 Education and self-rated health in Finland 18

4. The basis of self-rated health 19

4.1 Making health self-ratings 19

4.2 Subpopulations and the processes of health self-ratings 25

5. Previous studies on educational reporting heterogeneity 29

5.1 Self-rated health, education, and mortality 29

5.2 Diseases, biomarkers, functional health, and health behaviours 32

5.3 Summary and gaps in the evidence 36

6. Aims of the study 38

6.1 Research questions 38

6.2 Data 38

6.3 Sampling design 40

6.4 Final sample and missing data 41

6.5 Outcome variables 42

6.6 Health measures 43

6.6.1. Health behaviours 45

6.6.2. Clinical health 46

6.6.3. Functional capacity 49

6.6.4. Mental health and psychological well-being 51

6.6.5. Bodily sensations and pain 52

6.7 Statistical analysis 53

7. Results 58

7.1 Descriptive analysis 58

7.2 Logistic regression analysis 62

7.3 Interactions 64

7.4 Decomposition 67

8. Discussion 68

8.1 Comparisons with previous studies 68

8.2 Education as a confounder 71

8.3 Strengths of the study 74

8.3 Limitations 75

8.4 Conclusions and implications for future studies 80

Appendix 1. Abbreviations 108

Appendix 2. Literature search strategy. 109

Appendix 3. Clinical health variables 110

Appendix 4. Indicator correlation matrix. 112

HELSINGIN YLIOPISTO - HELSINGFORS UNIVERSITET - UNIVERSITY OF HELSINKI

Tiedekunta - Fakultet - Faculty	Laitos - Institution – Department
Faculty of Social Sciences	Department of Social Research

Tekijä - Författare - Author

Liina Marjukka Junna

Työn nimi - Arbetets titel - Title

Education and the basis of self-rated health.

Oppiaine - Läroämne - Subject

Sociology

Työn laji ja ohjaaja(t) - Arbetets art och handledare – Level and instructor	Aika - Datum - Month and year	Sivumäärä - Sidoantal - Number of pages
Master's thesis, Karri Silventoinen (University of Helsinki, Faculty of Social Sciences), Tuija Martelin (THL)	February 2016	83 + 5

Tiivistelmä - Referat – Abstract

Self-rated health (SRH) is a frequently used survey indicator of general health. It is periodically utilised in the study of educational health disparities. Several researchers have, however, suggested that systematic population sub group differences in health self-ratings (reporting heterogeneity) may results in SRH reflecting a different health status, or aspects of health, for different educational groups. Previous studies imply that the associations between SRH and other indicators of health may be strengthened by higher education. However, the studies disagree on the strength and the scope of the interaction effect. Comparability is also an issue due to, for example, the variation in the selected health indicators by which SRH is assessed. No such studies have so far been conducted in Norther Europe.

The purpose of this Master's thesis is to address educational SRH reporting heterogeneity. Using quantitative methods, this thesis analyses which aspects of health are included in dichotomised poor or very poor SRH ratings, and whether education moderates the relationship between SRH and the indicators of health. The selected health indicators represent five health dimensions identified in previous studies: clinical health, functional health, health behaviours, mental health and bodily symptoms and experiences. The analyses are conducted using logistic regression and regression –based nonlinear decomposition methods. The study utilises the Health 2000 data (n= 5586) for the household and institution dwelling population over the age of 30 residing in mainland Finland. The data is nationally representative and consists of a clinical- and mental health examination, and survey sections.

Overall, a high volume of somatic complaints was found strongly associated with poor self-rated health for all educational groups. Other significant contributors were functional health, diagnosed mental health conditions, and to some extent diagnosed diseases. An educational interaction effect was found for cardiovascular disease, subjective functional limitations in everyday tasks, and high volume of somatic complaints. In all cases education strengthened the association. However, for the majority of the indicators, SRH was associated with, no interaction effect was found. Compared to those respondents with a higher education, those with lower educational attainments more often reported poor SRH, but the selected health indicators and demographic variables explained virtually the whole difference. The study then, to some extent, concurs with earlier findings of higher education strengthening some of the associating between poor SRH and other indicators of health. However, the effect was statistically significant only when comparing basic education to higher educational attainments, and it was less systematic than some of the previous studies have suggested.

Avainsanat – Nyckelord - Keywords

Self-rated health, health status disparities, Finland, adult, logistic regression, education

Säilytyspaikka - Förvaringsställe - Where deposited

Muita tietoja - Övriga uppgifter - Additional information

HELSINGIN YLIOPISTO - HELSINGFORS UNIVERSITET - UNIVERSITY OF HELSINKI

Tiedekunta - Fakultet - Faculty	Laitos - Institution – Department	
Valtiotieteellinen tiedekunta	Sosiaalitieteiden laitos	

Tekijä - Författare - Author

Liina Marjukka Junna

Työn nimi - Arbetets titel - Title

Education and the basis of self-rated health.

Oppiaine - Läroämne - Subject

Sosiologia

Työn laji ja ohjaaja(t) - Arbetets art och handledare – Level and instructor	Aika - Datum - Month and year	Sivumäärä - Sidoantal - Number of pages
Pro gradu, Karri Silventoinen (Helsingin Yliopisto, Sosiaalitieteiden laitos), Tuija Martelin (THL).	Helmikuu 2016	83 + 5

Tiivistelmä - Referat – Abstract

Koettu terveys (self-rated health, SRH) on yksi kyselytutkimusten käytetyimmistä yleisen terveydentilan mittareista. Sitä käytetään usein terveyserojen tutkimuksessa. Aiemmat tutkimukset ovat kuitenkin kyseenalaistaneet indikaattorin vertailukelpoisuuden eri koulutustason suorittaneiden välillä, sillä ryhmät voivat painottaa erilaisia terveyden ulottuvuuksia tai mittapuuta vastauksissaan. Korkeampi koulutus voimistaa mahdollisesti koetun terveyden ja yksilön varsinaisen terveydentilan välistä suhdetta, kun terveydentilaa mitataan muilla vakiintuneilla indikaattoreilla. Tästä seuraa, että koettu terveys voi heijastaa erilaista todellista terveydentilaa eri koulutusasteen suorittaneilla.Tämän pro gradun tarkoitus on ensinnäkin selvittää, mitä terveyden eri ulottuvuuksia huonon koetun terveyden kokemukseen yhdistetään vertaamalla huonoa koettua terveyttä muihin vakiintuneisiin terveyden indikaattoreihin. Toiseksi, tarkoitus on tutkia koulutuksen vaikutusta koetun terveyden ja terveysindikaattoreiden väliseen yhteyteen.

Tutkimuksen aineisto on peräisin Terveys 2000 -poikkileikkaustutkimuksesta, joka sisältää terveyden indikaattoreita esimerkiksi lääkärintarkastuksesta, mielenterveyshaastattelusta ja kyselytutkimusosioista (n=5586). Valitut indikaattorit joihin koettua terveyttä verrataan edustavat kliinistä- funktionaalista- ja mielenterveyttä, sekä terveyskäyttäytymistä ja oireita kuten somaattisia vaivoja. Analyysit pohjaavat logistiseen regressioon ja epälineaariseen dekomponointiin.

Aineiston perusteella somaattisten oireiden suuri määrä on voimakkaasti yhteydessä huonoon koettuun terveyteen kaikissa koulutusryhmissä. Muita tärkeitä huonon terveyden selittäjiä ovat funktionaalinen terveys, mielenterveyshaastattelussa todetut mielenterveyden ongelmat ja osa lääkärin diagnosoimista sairauksista. Korkeakoulutus voimistaa huonon koetun terveyden ja sydän- verisuonitautien, koettujen funktionaalisten vaikeuksien, sekä korkean somaattisten oireiden määrän välistä yhteytä kun korkeakoulun käyneitä verrataan peruskoulutuksen suorittaneisiin. Koulutus ei vaikuttanut muiden terveyden indikaattorien ja huonon koetun terveyden väliseen yhteyteen.

Peruskoulutuksen suorittaneet kokivat useammin terveytensä huonoksi, mutta koetun terveyden ero selittyi pitkälti muilla terveyden indikaattoreilla mitatulla terveyserolla. Tutkimus on osittain ristiriidassa aiempien tulosten kanssa. Koulutuksen vaikutus oli oletetun suuntainen, mutta ero oli tilastollisesti merkittävä vain vertailtaessa peruskoulutusta korkeimpaan koulutusluokkaan. Koulutus ei myöskään vaikuttanut järjestelmällisesti kaikkiin terveyden indikaattoreihin.

Avainsanat – Nyckelord - Keywords

Koettu terveys, terveyserot, Suomi, aikuiset, logistinen regression, koulutus

Säilytyspaikka - Förvaringsställe - Where deposited

Muita tietoja - Övriga uppgifter - Additional information

1. Introduction

Self-rated health (SRH) is among the most frequently used survey indicators of general health (Jylhä, 2011). It is favoured for several reasons. It is a simple enough addition to any survey study and because of its simplicity, it is inexpensive to process and analyse. It has been continuously utilised in survey research since the 1950's. Therefore, an abundance of previous studies and data exists. Despite of its nonspecific nature or expressly because of it, a growing body of literature has confirmed SRH to be a good predictor for mortality. SRH has also been studied using other benchmarks for health than mortality - such as health behaviours, clinically measured biomarkers, functional limitations, or chronic conditions. The last decade has also witnessed a surge in studies dedicated to the processes of making subjective health assessments (Jylhä, 2011; Krause & Jay, 1994).

Self-rated health is, in addition to mortality, a frequently used health measure in the study of health differences between population subgroups such as socioeconomic-ethnic or age groups (Manderbacka, 1998a). Recent studies have, however, suggested that population subgroups may have systematically differing views on how SRH is assessed, or on which aspects of health should be included in the assessment (Dowd & Zajacova, 2010; Jürges, 2007; Smith, Shelley, & Dennerstein, 1994).

Previous studies have addressed this issue by, for example, comparing SRH's predictive power for mortality by population subgroups with mixed results. Testing whether the associations between health measures other than mortality and self-rated health are affected by education is still somewhat rare. Some previous studies suggest that "good" and "poor" health self-ratings do not necessarily correspond to an identical health status as indicated by more objective measures in all population groups.

Dowd and Zajacova (2010), Sohn (2015) and Schnittker (2009) found that good self-rated health corresponds to a different health status for those with a higher education, when compared to those with intermediate or basic education. Per Dowd and Zajacova (2010), when biological indicators of health are taken into consideration, those with a lower education and good self-rated health fare worse than those with a higher education at the equivalent SRH status. (Sohn, 2015) recently adduced that the associations between SRH, and health biomarkers, and -behaviours are strengthened by higher education. Both previously mentioned studies were conducted in the United States using the National Health and Nutrition Examination Survey (NHANES) data. This type of study has not been repeated in the context of the Nordic countries. Manderbacka (Manderbacka, 1998a) has studied the base of SRH in Finland in the 1990's but did not specifically focus on socioeconomic factors. Delpierre and companions (Delpierre et al., 2012) also suggest that Dowd and Zajacova's study be repeated in different contexts, while in a widely-cited article, Jylhä (2009) proposes that previously collected data be used for further studies on the basis of SRH.

The purpose of my Master's Thesis is therefore to take part in the discussion on the base of health self-ratings and combine it with the study of socioeconomic health differences. My aim is to analyse the relations between self-rated health and a wide variety of health indicators such as clinically measured illness, functional capacity, and biomarkers in Finland at the beginning of the 21st century. Furthermore, I will test if these relations vary by educational status.

This Master's Thesis in constructed as follows. Chapter 2 consists of a literature review on how health is defined and measured. The chapter provides the concepts used in this study, and introduces self-rated health. Chapter 3 reviews previous literature and recommendations concerning SRH and Chapter 4 the processes of making health self-ratings. Chapter 5 presents reporting heterogeneity. The research question, the data and methods used in the study are presented in chapter 6, and in the following chapter 7 the results. Chapter 8 is dedicated to connection and contribution to previous studies and discussion.

2. Measuring health

2.1 A working definition of health

Self-rated health has been called an indicator of general health status. However, what general health refers to is unclear. Health is not a specific term that a single, academically shared definition could cover (Jylhä, 2009). Neither is there a universal guideline for what is "good" or "bad" health, only a variety of indicators that point to a latent variable. In addition, what constitutes health has changed over time: over the last century, the definition of has expanded from measures directly relevant to survival and disease to include dimensions such as functional abilities, and quite recently also general well-being and quality of life (McDowell, 2006, p. 11).

One of the most well-known definitions of health is that presented by the World Health Organization (1948): "-- A state of complete physical, mental and social well-being and not merely the absence of disease or infirmity." In 1948, it summarised earlier attempts at positive definitions, and broadened the concept of health. It has later faced heavy criticism. If health must be this absolute, hardly anyone would ever meet the criteria of healthy. Disease patterns have also changed since 1948. The relative importance of acute infectious diseases has waned and current academics face the increase of degenerative and chronic conditions and diseases with interactive causes (Barsky, 1988; Blaxter, 1990, p. 4). A once useful opening has since been described as limiting and dated. At the beginning of the 21st century, the reality is the growing number of people living and aging with chronic disease. The WHO definition nevertheless works as a reminder of the importance of positive definitions of health on an idea level. (Huber et al., 2011; Karisto, 1984, p. 58-59.)

While the previous criticism targets the lack of pragmatisms of the WHO definition of health, authors such as Balog (2005) have challenged its multidimensionality. Balog notes that health is foremost a physical state and should not be mixed with concepts of well-being: "—health resides within individuals, it is natural phenomena, it is a state of physical wellbeing or physical fitness, and it is defined by how well the body is functioning in accordance with its natural design --" (Balog 2005, p. 270). While this view may be contested, on an analytical level it pays to demerge factors that *affect health, are affected by health,* and *health itself*.

Health is often defined negatively as the absence of ill health and disease, or positively, as WHO does, as enjoying good health and fitness (Von Wright, 1963). Manderbacka and colleagues (Manderbacka, Lahelma, & Martikainen, 1998) note that bad health often defines good health: bad health equals illness, while good health is actually an absence of the bad. Moreover, an illness is often thought to need a cause while a positive or normal state of health is not. For instance, in the late 1990's the Finnish people contributed illness and ill health to the following factors, in order of importance: stress and difficult living conditions, lack of physical exercise and bad diet and obesity (Helakorpi, Uutela, Prättälä, & Puska, 1999). According to Antonovsky (1987) health really should be seen as a continuum, while actual studies less frequently have a holistic or positive view on health (Idler & Benyamini, 1997).

In ecological definitions, health is defined as a balanced relationship between the individual as an organism and the surrounding environment. Instead of, or in addition to, looking at a disease as a reaction caused by the environment (biomedical view on illness), the focus is on the balance or imbalance of this interaction. When an organism, such as a human being, experiences physiological stress, it should be able to respond, adapt and return to a state of equilibrium. Failure to launch a response and to adapt would result in some form of physical strain or damage, which in turn may lead to illness (Coste, 2006). The view could be extended to include not just the biochemical and material environment, but also the social reality. This would include the possible challenges or even negative health effects of social adapting and the consequences and benefits of moulding the environment. (Karisto, 1984, p. 60.)

The dictionary definition sums up health as a "condition of being well or free from disease", thus combining the positive and negative definition. The medical dictionary goes into more detail, describing health as: "1. the condition of an organism or one of its parts in which it performs its vital functions normally or properly, the state of being sound in body or mind, freedom from physical disease and pain" and "2. the condition of an organism with respect to the performance of its vital functions especially as evaluated subjectively or nonprofessionally" (Merrian-Webster Dictionary, 2016). This definition is already more inclusive, since it adds the subjective experience, bodily experience of pain and functional health.

Following this, health appears multi-dimensional. It includes aspects such as psychological health, disease and organ functions, and pain. The multidimensionality is also often reflected in lay talk of health. Several researchers have noted this, and attempted to identify the references most often used in lay (or medical professionals') definitions of health. Most of these studies classifying health factors that laypeople include in their health assessments seem to agree on the existence of at least three, widely referred to dimensions: biological/medical (measures based on physiological ideas of normality), social interactional/functional (health in terms of being able to perform normally in social roles; both in the view of the subject and of physical incapacity) and the subjective model (subjective experiences).

As an example, one qualitative interview study (Kaplan & Baron-Epel, 2003) identified three models by which the study participants judged their health status. These models were the biomedical or disease oriented model, the emotional or general feeling -model and the functioning-related model. Ware (1986) summarised the preceding studies classifying health dimensions into physical-, mental-, social-, role- and general health perceptions and symptoms. Manderbacka's (1998) health model consisted of absence of ill-health (disease states), health as function (a non-restrictive state), health as experience (bodily and mental, not defined medically), health as action (result of health behaviour) and other (normal function of senses, sexuality & reproduction). Borawski and colleagues (Borawski, Kinney, & Kahana, 1996) also grouped the health musings of elderly survey respondents into five categories: physical health, health-transcended, attitudinal/behavioural health, the externally focused and the non-reflective health. Blaxter (1990, p. 17-34) saw physical fitness, general energy levels, health as a resource and health as a factor in social relationships as other possible dimensions. In one qualitative study (Idler, Hudson, & Leventhal, 1999), the final number of dimensions was six: physical health (diagnoses and symptoms), physical functioning, health risk behaviours, social role activities, social relationships, and the psychological, emotional, and spiritual health.

In this study, health is defined holistically and very broadly. Following the previous suggestions, health is understood to primarily consist of medical health (absence or presence of disease, health biomarkers) and mental health, but it also includes health behaviours (associated with future health status, health as doing), functional health and unspecific bodily experiences.

2.2 How to measure health

While there may not be a single operational definition for health, there is a constant need for reliable and cost-effective methods for measuring it. Health is of interest to medical professionals, politicians, policymakers, researchers, employers, and economics, as well as the laypeople. Health measures are utilised in individual health care as well as in public health, for example when measuring changes in population health or identifying vulnerable persons and populations. The question is, how best to measure something that is difficult to even define.

Since no simple measure or even a combination of several measures covers the full complexity of health status, its indicators are always chosen and purposefully developed. The underlying concepts and definitions, along with social norms and concerns outline how health is defined, categorised and measured. Health measurements are often categorised as either subjective or objective (also known as internal and external, respectively). Both can measure health status and its change for either an individual or for a population. (McDowell, 2006).

Even though many academic disciplines study health, in the public health tradition whether a person – or a population for that matter – is healthy, is ultimately evaluated by clinical standards. *Objective health measures* are based on the biomedical definition of health and illness presented in the previous chapter. Normalities and abnormalities of the body and its functions are assessed following the prevailing clinical standards. This is done by measuring and analysing some observable attribute, state or symptom by the current categorised pathologies and norms. A healthy person is one free of observable pathologies (Karisto, 1984, p. 66).

As Blaxter (1990, p. 2-3) and Karisto (1984, p. 66) note, even these observable pathologies, which are often thought of as objective indicators of health, are to some level socially constructed. The functioning of organs is not a binary matter of healthy/unhealthy, but a continuum. Laboratory results of an indicator may vary even between individuals that are all deemed healthy. Using the current medical standards, it may still sometimes be difficult to draw a line between normal and abnormal. Clinician's do not always agree in their verdicts, nor are they always able to find a matching explanation for symptoms. Many illnesses are caused by several factors, such as life style and genetics, with multiplied effects (Huttunen, 2012). Furthermore, clinical health evaluations conducted by professionals are not altogether free of the subjective component either, and should not be treated as such. Which tests in practice are conducted for the patient and which results are deemed relevant are matters of choice on the part of the clinician (Bjorner et al., 1996), while a recent literature review found the doctor-patient relationship to some degree influenced by the patient's socioeconomic background (Willems, De Maesschalck, Deveugele, Derese, & De Maeseneer, 2005).

An alternative to objective measures are the *subjective health measures*. They collect information on the health experience of the subject. Self-rated health (SRH) is one such measure. The type and range of subjective health information deemed relevant and valid has changed considerably over the past decades. Until the 1970's, health surveys mainly defined health as the absence of clinical disease. Thus, when subjective measures of health were collected, they often functioned as a proxy for unavailable clinical health measures (Manderbacka, 1998a). The study participants were often for instance asked to report if they had diagnosed diseases or other medical conditions. The underlying assumptions behind these measures therefore alluded to the medical definition of health and the superiority of objective measures. The respondents were expected to a reasonable degree report conditions that a clinician would confirm. Since then, the definition of health in social research has broadened. Self-ratings of symptoms, bodily experiences, health behaviours, self-reported diseases, functional ability, health-related quality of life and psychological well-being are all examples of subjective health measures (Bjorner et al., 1996).

Both the subjective and objective indicators have their benefits and limitations; both are used widely for measuring health, by a variety of stakeholders. Objective health states, such as illnesses and pathologies, are associated with subjective health indicators (Goldman, Glei, & Chang, 2004; Manderbacka, 1998a), but subjective health is also affected by other factors (Jylhä, 2009), such as knowledge of medical standards (Goldman et al., 2004), the bodily experience and medical history (Benyamini, Leventhal, & Leventhal, 2003; Idler, Leventhal, McLaughlin, & Leventhal, 2004), relative wellbeing and knowledge and experiences related to aging (Blaxter 1990, p. 3).

Measuring objective clinical health seldom needs to be explicitly justified. Whether an individual has or does not have a medical condition is usually seen as unquestionably valuable information. Accordingly, so is a population level change of an objective indicator of health. Automatically prioritising the objective health data over the subjective has however been criticized from a variety of viewpoints. It has been suggested that focusing on the objective is insensitive to an individual's own experience and understanding of health. Also, some aspects of health simply cannot be objectively measured. For instance, pain is a sensory experience and therefore only observable by the person experiencing it. Some academics have suggested that individual bodily experiences, while difficult to measure directly, may be valuable precursors for pre-clinical state health conditions (Jylhä, 2009; Sen, 2002). Majority of the studies that promote the importance of self-rated health base their arguments on the uniqueness of the health information on current and past health held by the individual (Idler, Russell, & Davis, 2000; Jylhä, 2011).

The recent hype around future health care technologies and personalised health has also brought the focus on subjective experience by way of questions of ownership of health (Boccia, 2012). For an example, in the UK the 2001 Department of Health doctor-patient relationships guideline prompts clinicians to meet the patients as experts of their own health and to promote their knowledge, participation, and self-care (Kennedy, Rogers, & Gately, 2005).

The importance of the experience of health can then be promoted either as in relation to clinical health or by itself, as lived-in experience of wellness. When advocating the *subjective view*, the individual is considered the owner of the best possible information, even up to a point where the individual should have the right to choose the criteria by which health is defined (Karisto, 1984, p. 14-17). This approach can be questioned taking the *objective view,* per which health is a biological state, and not a matter of opinion. A medical professional will follow a scientific procedure and gather evidence-based data followed by an educated judgement. Thus, individual judgements measure something other than health. After all, individuals are often unaware of both their actual health state, and factors that promote or hinder health processes.

As an epistemological question, if reliable and useful health information can be acquired from the individual, does health still equal objective health, to which the subjective is simply a window (Karisto, 1984, p. 22-26) or is the subjective valuable in itself? And yet, as Sen (2002, p. 861) states:

> *"Although the internal view is privileged with respect to some information (particularly that of a sensory nature), it can be deeply deficient in other ways. There is a strong need for scrutinising the statistics on self-perception of illness in a social context by taking note of levels of education, availability of health facilities, and public information on illness and remedy".*

Subjective health self-ratings such as the SRH are made in relation to the historical context, the cultural definitions of health, the socioeconomic and status-related expectations of health possibilities and risks, as well as the individual medical histories and personality characteristics (Etilé & Milcent, 2006). Also, pre-existing worries about health as well as other sources of stress have been suspected to play a part in some self-reported health measures (Fylkesnes & Førde, 1992). Measures of hypochondria have been important for the subjective health of medical outpatients, along with somatic symptoms and disability (Barsky, Cleary, & Klerman, 1992).

Perhaps because of this complexity, mortality is an often-used indicator as the Gordian Knot -solution to measuring health. Death is indeed “the final objective state of health”, as Quesnel–Vallée (2007) put it, since a person who dies as a cause of a disease can safely be called less healthy than a person who does not. Still, at the beginning of the 21st century, aging even for decades with a medical condition is the reality for a growing portion of the population, so years lived healthy and the experience of health need to be studied further (Blaxter, 1990).

2.3 Education as a determinant of health

Health may be a multidimensional physical state, but it is also socially patterned. That inequities in social and physical environments lead to inequalities in health was already observed in the 19th century. According to Blaxter (1990, p. 6) the academic interest in health inequalities spiked in the late 1970’s. Throughout and after the decade, several studies addressed the socioeconomic and educational differences in health (for example, see Bunker, Gomby, & Kehrer, 1989; Myers, 1974; Townsend & Davidson, 1982) using indicators such as mortality and subjective health assessments. In Finland, health inequalities have been studied since the 19th century (Karisto, 1981). One distinctive feature of the Finnish study tradition is the use of the high quality national statistical and register data and population health surveys, both for the purposes of descriptive studies of health inequalities in Finland, as well as the underlying patterns and causes (Palosuo et al., 2009).

By 2016, the relationship between education and health has been well documented (Conti, Heckman, & Urzua, 2010; Mackenbach et al., 2008), and educational health disparities are also currently being monitored in several countries. Education has been used in the study of health disparities both as an independent coefficient and as part of the concept of socio-economic status.

As Shavers (2007) summarises, education is associated with several health conditions, behaviours, and attitudes, as well as factors relevant to health, such as living and working conditions and material environments. It is largely agreed upon that an educational divide still exists even in the wealthy Europe. Those in lower educational positions have consistently had higher mortality rates, lower self-reported health and higher rates of chronic conditions and disease (Mackenbach, 2006).

While the education-health -association is well known, the social and biological processes behind it are neither fully understood, nor fixed. However voluminous the literature on the subject, whether these disparities are due to education as such, or some other relevant factor such as material living conditions is still being contested. It is also unclear if education affects the health of entire populations in a uniform manner, across for instance sex, race, and age groups (Shavers, 2007). Most often cited explanations for the association are *causal mechanisms (*either due to behavioural, or material factors), *selection* (reversed causality) and *artefactual mechanisms*.

The first explanation refers to a *causal association* between education and health status. Since education cannot provide protection from illness like a vaccination against disease, the association is likely to be indirect. Causal explanations are often separated into the material and the cultural. Material benefits such as better living conditions and access to quality health care often result from higher educational attainments. Another study tradition of causal explanations emphasises cultural factors, such as educational group health behaviours, which may over time lead to differences in health. The *selection explanation* on the other hand states, that the educational status of a person may in fact be caused by the health status, not vice versa. Someone who is often ill or has a chronic condition is less likely to succeed academically. The last explanation, the *artefact*, refers to mechanisms of measurement error and study limitations. The latter explanation is seldom thought to be the sole factor behind the observed educational differences in health; rather the differences in health are caused by complex causal mechanisms. (Cutler & Lleras-Muney, 2006; Goldman, 2001.)

3. Self-rated health (SRH)

3.1 Definition and recommendations

As the previous chapter illustrated, various stakeholders have an interest in measuring health, and there are also various options available for the purpose. Survey single question self-rated health is one such option. Since objective health data measured in a medical check-up is expensive and time-consuming to collect, enquiring the subject for a self-assessment may be economical (Etilé & Milcent, 2006). Health is thus a standard theme in survey research. Several questionnaires and survey items have been developed for the purpose.

As an example, the SF-36 36-items questionnaire tool for health self-assessment targets the broad definition of health by individually addressing eight specific dimensions of health: limitations in physical activities because of health problems, limitations in social activities because of physical or emotional problems, limitations in usual role activities because of physical health problems, bodily pain, general mental health, limitations in usual role activities because of emotional problems, vitality and general health perceptions. These eight scales are utilized with equal weights to build a health scale from 0 to 100. (Ware Jr & Sherbourne, 1992.)

Single item *self-rated health* (SRH, sometimes called self-assessed health, subjective health, or perceived health) is a widely-used alternative to such multiple item health questionnaires and is also included as an item in the SF-36. SRH is often used in large scale surveys. It is a subjective measure of general health that is typically structured as a Likert Scale. The answering choices available may vary, but the most commonly used alternative is a five-point scale formulated as some variation of the following phrasing: “Would you describe your health as very good, good, fair, poor or very poor?” (*Global self-rated health*). The question can be more detailed. For an example SRH may be formulated to inquire the subject's health in comparison to his views on the health of his peers (*comparative self-rated health*). The question may also involve a time dimension, either inquiring health at a certain time or compared to a time period.

The World Health Organization (World Health Organization, 1996) and the European Union Commission (COM(95), 1995) recommend the tool as a standard inclusion to health surveys. The action on European Community Health Indicators (ECHI) has listed SRH as one of its 88 comparable key health indicators for European health research (Verschuuren et al., 2013) and the measure is also a recommended for measuring disparities in OECD countries in an OECD Committee Health Working Paper (De Looper &

Lafortune, 2009) and for the European region by WHO (De Bruin, 1996). Major national surveys such as the World Value Survey, European Value Survey, the National Health, and Nutrition Examination Survey (USA) and SHARE (Europe) all gather SRH measures regularly.

The World Health Organization (Subramanian, Huijts, & Avendano, 2010), the European Union Statistics on Income and Living Conditions EU-SILC (Eurostat, 2016), European Health Interview Survey EHIS (EHIS, 2010) and the European Social Survey (ESS, 2016) formulate SRH as follows: “In general, how would you rate your health today?”, with the choices: “very good”, “good”, “moderate/fair”, “bad” to “very bad”. For instance, the Survey of Health, Ageing and Retirement in Europe (SHARE, 2016), the National Health and Nutrition Examination Survey NHANES conducted in the United States (Idler & Angel, 1990) and the Joint Canada/United States Survey of Health all have chosen the wording: “Would you say your health is: Excellent, Very good, Good, Fair, Poor”. For the sakes of clarity, the aforementioned wording will from now on in this thesis be addressed as SAH or the asymmetrical wording to separate it from the EHIS and ESS recommended wording. This is not a conventional practice in the study of subjective health measures, but since the wording appears to be somewhat more usual in North American research literature and the wording is regularly used under the title of self-assessed health in that area, this choice is convenient.

It is still unclear if the various wordings of self-rated health questions are interchangeable. This study will focus on the symmetric EHIS and ESS formulations in which the neutral option of the 5-point Likert scale is the middle one. For the sake of clarity, the formulation more often used in North America with three positive option choices, one neutral and one negative (presented above) will from now on in this study be called self-assessed health, SAH.

The benefits of the self-rated health measure are numerous. Compared to multi-item measures SRH is less burdensome to collect and therefore an inexpensive addition to a survey. It is interpretable as it is. For instance, in the case of the SF-36 tool, the purchase of commercial software is necessary for interpreting the scores. Other benefits of SRH include its reliability, reproducibility and validity and its verified correlation with many other indicators of health, both objective (such as biomarkers and physician ratings) and subjective (multiple item health indicators) (DeSalvo et al. 2006; Jarczok et al. 2015). The indicator's implicit, general, and open definition of health itself can also be an asset. Bjorner and colleagues (Bjorner et al., 1996, p. 7-8) state that the measure can be used as a health outcome in itself, or as a predictor for desirable or undesirable health outcomes, for instance when identifying high risk population sub groups. SRH is a predictor for mortality (Jylha 2009) and a variety of health outcomes, and therefore useful as a risk-assessing tool in both clinical (DeSalvo et al. 2006) and health policy environments (Bierman et al. 1999). It can also be utilised as a measure of general health along with more specific measures when evaluating clinical interventions (Bjorner et al., 1996, p. 9).

3.2 Psychometry

A good measure comprehensively covers the phenomenon that it targets, produces reasonably congruent results with other measures of the same phenomenon and is a good predictor for some outcome. In other words, a measure needs to be valid, reliable and its sensibility to change must be known. Besides this, a measure needs to be feasible and useful in practice in terms of costs and the trouble of collecting it. Whether this condition is met can also be evaluated by the people utilising the measure. In the case of SRH, the free use of the measure and its interpretability are both benefits. (De Vet, Terwee, Mokkink, & Knol, 2011.)

As a tool for measuring health, SRH has been exposed to a substantial amount of study. Survey measures like SRH collected by standardised data collection procedures is most often used to generalise estimates from a sample to a larger frame population. Like any data collection method, the survey has its benefits, limitations, and underlying assumptions. The standardization of questions, for example, is done to improve statistical reliability, while the artificial response alternatives, such as from very good to very poor, have been criticized for diluting validity. Validity may also be compromised if the respondents interpret the question or the survey situation in some way that was not intended by the researchers (Manderbacka, 1998a 2).

Validity consists of several dimensions of which those essential for the use of the indicator must be considered. *Simple face validity* questions whether the measure is intuitively acceptable and relevant to users. Jylhä (2009) notes that the question is understandable to respondents and feasible and convenient in clinical trials, general practice, and risk assessments. A measure with good *content-based validity* will cover the entire target phenomenon. Several studies have suggested that SRH is associated with a wide range of health dimensions as well as mortality and morbidity (Bjorner et al., 1996, p. 34-35; Idler & Benyamini, 1997; Jylhä, 2009; Sohn, 2015).

C*riterion-based validity* consists of two dimensions. The first one is *concurrent validity*: the correlation between the measure and some accepted measure of the same phenomenon. The second dimension is *predictive validity*: the correlation between the measure and some accepted measure of the phenomenon in the future. Several statistical tests exist to assess validity (De Vet et al., 2011). The problem with general health is that no golden standard for measuring it exists, and therefore conducting tests of for example content and criterion-based validity, is challenging. Consequently, some measure that is thought to best represent health should be selected to be able to test SRH. Never mind the issues mentioned in the previous chapter, a full clinical examination and disease-related mortality are still often used as the superior and criteria for this purpose. (Bjorner et al., 1996 p. 34-35; Kaplan, Bush, & Berry, 1976.)

SRH has been tested both for concurrent and predictive validity in numerous studies. Several of these studies compare SRH measures to multi-item survey tools (Lindeboom & Van Doorslaer, 2004; Rowan, 1994) or objective indicators such as biomarkers, reaction time or stimulus, to test *concurrent* validity. In other words, the indicator's relation to some current health indicators. Mortality would be an example of SRH's *predictive validity*: the SRH measure is collected at one time and compared to later outcomes (Idler et al., 2000; McDowell, 2006 31-32).

Reliability assesses the stability and consistency of the measure as well as its inherent measurement error. In the case of health measures, this usually means testing either repeatability or reproducibility. Whether a measure is stable over a period should the health status of the person remain unchanged (test-retest in similar conditions) is filed under repeatability. Reproducibility on the other hand targets the variation between tests conducted in diverging environments; in the case of the SRH for instance between two countries or two points in time.

While a cross-sectional study setting is most often used in studies that address SRH, longitudinal panel studies have also been conducted to assess the stability of the measure over a period of several years. Johnson and companions (Johnson, Stallones, Garrity, & Marx, 1990) and (Rodin & McAvay, 1992) assessed SRH and its relation with associated variables over periods of one and three years, respectively. The previously mentioned study found several health indicators, such as increase in physician visits and new illnesses and negative developments in pre-existing conditions, to be associated with a decline in SRH. However, other factors such as baseline depression and low life satisfaction were also associated with the decline. A continuing decline in SRH was found to be a determinant for increased future depression.

Reliability tests may also measure whether different wordings or versions of the measure can be considered identical in contents. However, testing this over different wording or scaling options of global SRH is very rare. This is possibly because the study setting would require asking the subjects to rate their health several times in several ways, which would likely require gathering new data specifically for this purpose. It is plausible that SRH measures are not entirely comparable between different wordings. More often than testing the wordings, comparisons have been made between global and comparative self-rated health (Baron-Epel & Kaplan, 2001; Vuorisalmi, Lintonen, & Jylhä, 2005; Vuorisalmi, Lintonen, & Jylhä, 2006).

Some studies have treated the wordings as interchangeable. For instance, when Johnson and companions (Johnson et.al. 1991) intended to analyse whether health self-ratings are summary indicators for physical and psychological health or a proxy for physical health, they analysed SRH’s association with several variables using four different samples with varying wordings. The researchers utilised the data sets to assess if the associations remained the same between the samples as if the SRH measures were identical. Erikson and companions (Eriksson, Unden, & Elofsson, 2001) on the other hand purposefully compared various determinants of self-rated health such as demographic, physical, functional, and mental health to two global SRH wordings (scales from 1 to 5 and 1 to 7) and an age comparative wording. Overall, the correlation structures of the three self-rated health measures were found similar, while the strength of the correlation varied slightly from item to item. The 7-point non-specific response option SRH scale was found to lead to more missing data in the case of elderly participants (Bjorner et al. 1996, p. 34-35, p. 49).

Other issues included in psychometric evaluations are sensitivity to change and interpretations and cross-cultural applicability, which may be used to argue validity as well, even though these aspects are often presented separately from validity. Sensitivity to change is especially important when a measure is used to evaluate interventions: does the measure capture change and lack thereof reasonably well? According to an often-cited study (Bailis, Segall, & Chipperfield, 2003), SRH change followed change in physical and mental health and to some extent that in social support and health-related behaviours. In the study both physical- and mental health were self-reported, but many studies do suggest health self-ratings to be a dynamic evaluation (see K. F. Ferraro & Kelley-Moore, 2001; Stenholm et al., 2016); for instance, death has been shown to be preceded by a decline in self-rated health. (Bjorner et al. 1996, p. 35-36.)

Self-rated health is vague. It offers very little definition of how health should be interpreted and in relation to what or whom. Therefore, different individuals may incorporate different elements to the question. Some may define health strictly by their understanding of physical health, while others may have a more holistic view that includes some aspect of psychological or emotional well-being. As a general rule of survey technique, imprecise questions are likely to have low reliability (Rossi, Wright, and Anderson 2013). However, SRH has in various studies been found to be both reliable (test-retest) and valid. The results it produces are in line with many other health measures (DeSalvo et al. 2006; Krause & Jay 1994).

Several studies have also found that even after controlling for various other indicators of health status (both objective and subjective) and behavioural, psychological, social and environmental factors SRH still predicts health outcomes, such as mortality (Idler

& Benyamini, 1997; Idler et al., 2000). SRH's test-retest reliability has also been compared with that of other ways of inquiring about subjective health - for instance by posing very specific health questions related to the symptoms of an illness – with the conclusion of good reliability (Lundberg & Manderbacka, 1996).

3.3 Education and self-rated health in Finland

Large scale, nationally representative population health surveys have been conducted regularly in Finland since the 1960's (Prättälä et al., 2007, p. 15). Self-rated health has been included regularly in Finnish surveys such as the Mini-Finland Health Survey (Aromaa & Koskinen, 2004), and Health Behaviour and Health among the Finnish Adult Population (Helakorpi et al., 1999) since the 1970's.

Since the 1970's, the Finnish standard of living, working conditions and health knowledge have improved while ischemic heart disease and stroke mortality have been severely reduced. Between the 1970's and 2004, the overall self-rated health of the Finnish population also improved dramatically among both sexes and in all age groups (Heistaro, Vartiainen, & Puska, 1996; Helakorpi et al., 1999; Rahkonen et al., 2004). And yet, the magnitude of the change has been greater for some population groups; those with a higher income and/or education.

Despite the population-level positive developments, several studies have concluded that the state of Finnish population health is undergoing two simultaneous, contrary processes. There is a documented dissonance between the improved general population health, and the increase in health inequity (Lallukka, Rahkonen, Lahelma, & Laaksonen, 2011). This educational and occupational class inequity is especially notable in the light of Finland's small income inequality. Education along with household income (Heistaro et al., 1996) has been found associated with SRH. At the beginning of the 21st century, a separate trend remained for the educational groups: with high education came better health. The overall population level differences by education diminished slightly during the 1990's, but for the women the difference actually grew again between 2002 and 2004. Heistaro and colleagues (Heistaro et al., 1996) studied SRH, education and income between 1972 and 1992, after a period of economic growth had just turned into a severe depression, while Rahkonen and colleagues (2007) addressed the period between 1980 and 2004 using education and occupational class.

Both improving the population health and decreasing the inequity between population sub groups has been on the target list of Finnish health policy since the 1980's (STM, 1986; STM, 2001), only so far it appears that the former objective has had more success (Mackenbach et al., 2008).

4. The basis of self-rated health

4.1 Making health self-ratings

In Chapter 3, self-rated health was presented as one measure of general health. But since self-rated health does not include a definition of how health should be assessed, the question is what frame of reference the respondents use and how conclusions are reached based on symptoms and illness (Bjornen et al. 1996, p. 23-25).

Unless stated otherwise, all the studies included in this Chapter utilised a 5-point Likert scale global self-rated health. Because of the variation in the wordings of the question utilised in previous studies, I have chosen to also to include other wordings as the one recommended in Chapter 3.1, as well as studies that have dichotomised the answers. The early analyses of self-rated health seldom problematized or reflected on the specific wording of the question, but the studies are nevertheless relevant as a starting-point. Studies addressing time-comparative SRH were excluded. Asymmetric global self-assessed health (SAH) is separated from SRH.

The social and biological base of self-rated health has been studied extensively since the mid-20th century, and yet most studies rush to state we still do not fully understand the processes of health self-rating. Previous studies have confirmed that age, sex, several indicators of health behaviours, self-reported chronic conditions and diagnosed clinical diseases all affect health self-ratings. Per several researchers (Blaxter, 1990, 13-14; Kaplan & Baron-Epel, 2003), the biomedical model of health deeply affects lay health assessments. Still, some of the factors related to SRH are more and some less related to medically defined health. The following theoretical model (figure

1) is based on the suggestions of previous studies (Fylkesnes & Førde, 1992; Jylhä, Leskinen, Alanen, Leskinen, & Heikkinen, 1986; Moum, 1992).

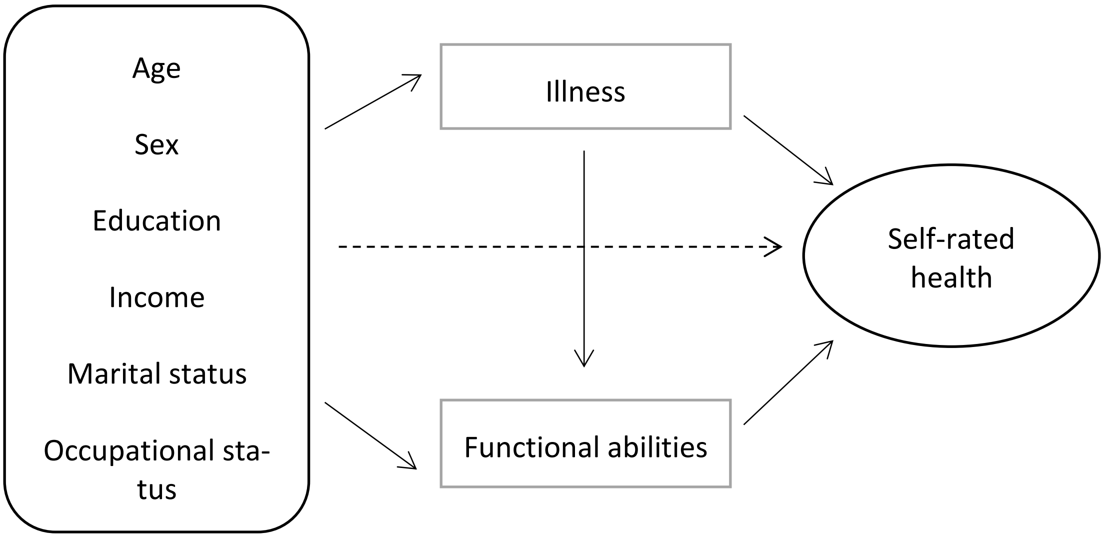


Figure 1. Theoretical determinants of self-rated health.

Following Chapters 2.2 and 3.2, one way of analysing what self-rated health consist of is to test it against some other health measure using statistical methods. Already in the late 1950’s, Suchman and colleagues (Suchman, Phillips, & Streib, 1958) noted that the self-assessments of health are correlated with clinical ratings, but seem to include other information also. In the classical study, the researchers compared clinician- and self-ratings of health made on an identical 5-point scale. Additionally, four self- and physician -reported conditions were compared: impaired hearing, varicose veins, haemorrhoids, and abnormalities of the scrotum area. In the second part of their analysis, SRH and its relationship with factors that the researchers called health behaviours and attitudes is analysed. However, in the study health behaviours mostly consist of questions such as: “Has your health or physical condition ever interfered in any way with your ability to do your job?”. The study is a fantastic opening statement, but whether the doctor’s assessment truncated into a value between 1 and 5 can be feasibly called “actual health” as the study does is debatable. The same holds for matching judgements on the prevalence of the selected health conditions.

Maddox (1962) found self-rated health as measured by a 4-point Likert scale to be in line with the results of medical examinations in the small sample of 270 elderly participants. Later, in a 15-year follow up of the same study Maddox and Douglas (1973) followed with an often-cited observation: self-ratings simultaneously measure something more and then again less than their objective counterparts.

The latter study (Maddox & Douglass, 1973) has in numerous occasions been called a classic on subjective health. The association between a clinician's evaluation and self-rated health remained throughout the study period, although the strength of the association was only moderate. Self-rated health at the starting point proved to better predict self-rated health 15 years later, while the doctor's assessments were found poor predictors of clinical assessments 15 years later. In addition, self-rated health measured in the first study phase predicted clinical assessments in the last phase better than vice versa. When the subjective evaluation deviated from the clinician's, the subjective tended to lean towards optimistic. The study used data from six survey collection points, with varying subjective health measures: health is sometimes inquired using a 4-point scale and sometimes as comparative self-rated health, as in relation to the participant's age.

Major contributors to global, symmetric 5-point SRH assessments are medically defined health and the related impact on the functional abilities and bodily experiences of the respondent (Eriksson et al., 2001; Moum, 1992). Several other studies have produced similar results using other subjective health measures such as other wordings or scales (Wolinsky, Callahan, Fitzgerald, & Johnson, 1993).

To my knowledge, out of objectively measured biomedical outcomes hip fracture for women (Cummings et al., 1995), low-grade inflammation for young men (Warnoff et al., 2016), impaired glucose tolerance for men (Andersson et al., 2013) functional ADL disability (Kaplan, Strawbridge, Camacho, & Cohen, 1993), decline in physical functioning and functional limitations (Idler & Kasl, 1995; Idler et al., 2000), motor skills related speed functions (Era, Jokela, & Heikkinen, 1986), and functional ability as rated with scales (Schultz-Larsen, Avlund, & Kreiner, 1992), have been found associated with 5-point global SRH. The asymmetric 5-point SRH has also been studied in relation to cause specific mortality and while it was found a good predictor for all-cause mortality, it predicted specific causes of death such as diabetes and respiratory and infectious diseases particularly well (Benjamins, Hummer, Eberstein, & Nam, 2004). Several self-reported health outcomes such as hypertensions, myocardial infarction, asthma, heart failure and stroke (Chan et al., 2015; Tibblin, Cato, & Svardsudd, 1990) are also associated with SRH.

Not all health conditions however influence subjective health assessments equally, or even directly. Subjectively experienced physical condition and performance have been identified as important for self-ratings. Chronic conditions and disease that directly affect this physical condition are also associated with the measure (Fylkesnes & Førde, 1991; Goldstein, Siegel, & Boyer, 1984; Pope, 1988). However, symptomless medical conditions or conditions with long symptomless periods are less often taken into consideration in making health self-ratings (Manderbacka, 1998). A preliminary study (Jorgensen, Langhammer, Krokstad, & Forsmo, 2015) concluded, that whether the person is aware of having a disease or not is a significant factor in making health assessments.

Biomarkers are measurable indicators of physical processes. Changes in certain biomarkers may imply an increased in vulnerability towards disease processes. Several researchers nominate the relationship between SRH and biomarkers a promising future area for research (Dowd & Zajacova, 2010; Fayers, 2005; Jylhä, 2009). Biomarkers such as indicators of inflammation, HDL ratio, cholesterol levels, haemoglobin levels, the ratio of cortisol to DHEA-S, blood pressure, and BMI (non-linear association) (Christian et al., 2011; Goldman et al., 2004; Wu et al., 2013) are associated with SRH, although the findings concerning BMI are not fully consistent (Chan et al., 2015). Poor SRH in males is associated with increased s-prolactin and decreased s-testosterone levels, which in turn have both been identified as biological response to prolonged experience of stress (Halford, Anderzén, & Arnetz, 2003).

Also, factors other than clinical illnesses and indicators of biological processes affect self-ratings. Health behaviours, such as following diet recommendations for women

(Collins, Young, & Hodge, 2008), and physical activity, smoking and alcohol consumption for both sexes are also associated with self-rated health (Fylkesnes & Førde, 1992; Idler & Angel, 1990; Szwarcwald et al., 2015). Health-related symptoms and bodily experiences like chronic pain (Mäntyselkä, 2004) as well as psycho-social well-being and health are also relevant factors (Blaxter, 1989; Farmer & Ferraro, 1997; Jylhä et al., 1986; Manderbacka, 1998b). Such bodily experiences as fatigue and diminished capacity for activity (Benyamini et al. 1999), subjective feelings of vigour, objective functional capacity and fitness and subjective tiredness (Shirom et al. 2008) are associated with changes in SRH.

The association between health behaviours and self-rated health has not been systematically confirmed across study settings. Manderbacka and colleagues (Manderbacka, Lundberg, & Martikainen, 1999) analysed the association between a 3-point symmetric SRH scale and various self-reported health behaviours and body mass index in Swedish adults. At first glance, vitamin use, diet, obesity, being underweight, smoking and taking part in physical exercise were all found associated with the self-rated health measure. But after the data was adjusted for self-report scales of physical and mental health problems and functional limitations, only smoking and vegetables in diet retained their association in the whole population and obesity and underweight for young adults. Hereby the association between most health behaviours seems mediated by negative health outcomes. SRH seems, to a larger extent, reflect actual health problems, than potential ones.

Bodily sensations may function as indicators of underlying physiological processes (Idler et al., 2004; Jylhä, 2009). Since these bodily sensations are available only to the person experiencing them, this information is known to the subject alone. Some preclinical stage diseases may thus imperceptibly affect SRH (Jylhä 2011). A feeling of, for instance, general lack of energy does not equal a clinical condition, but cumulative bodily experiences may inform the individual of ongoing health processes.

Bailis and colleagues (Bailis et al., 2003) note that SRH is *spontaneous* as in responsive to current, observable illness; simultaneously it is *enduring* and correlated with future self-assessments of health and functional abilities. Several studies (Heistaro, Jousilahti, Lahelma, Vartiainen, & Puska, 2001; Manderbacka, 1998) have observed that chronic conditions, however, have a stronger association with SRH than acute illnesses. In Manderbacka's (1998) interview study respondents differentiated between temporary impairments to health such as bronchitis, and the type of illnesses and temporary conditions such as the flu that are considered "normal" and as such not a threat to general health. Also, certain symptoms and conditions were categorised as a part of either the aging process or for instance caused by the respondent's working conditions and therefore considered separately from other medical states that were viewed as long-standing illnesses. SRH assessments based on the current health status are also fluid and open to change (Ferraro & Kelley-Moore, 2001).

SRH is also a predictor of future chronic conditions (Shadbolt, 1997), health service needs (Smith, Glazier, & Sibley, 2010), future use of medical services (Bath, 2003) entering assisted living for the elderly (Weinberger et al., 1986) and the risk of exiting the work force (Lund & Borg, 1999). Furthermore, it predicts mortality from specific diseases such as coronary heart disease (Bosworth et al., 1999), cardiovascular diseases (Stenholm et al., 2016), and cancer (Shadbolt, Barresi, & Craft, 2002). These findings strengthen the importance of subjective health assessments for identifying pre-clinical conditions.

Good health is often defined as lack of illness. However, Kaplan and Baron-Epel (2003) suggest that positive and negative subjective health assessments are not symmetrical: different factors influence them at the opposite ends of the scale. Good health is a passive state that simply happens, while bad health -as in occurrence of disease- requires a cause (Kaplan & Baron-Epel, 2003). Some studies hence suggest that the determinants of good health are not a mirror image of those of poor health (Mackenbach, van den Bos, Joung, van de Mheen, & Stronks, 1994; Smith et al., 1994): a healthy lifestyle and good functional capabilities would be foremost associated with better SRH, while mobility limitations, pain, depression and stress affect SRH negatively (Delpierre et al. 2009; Goldman et al. 2004; Jarczok et al. 2015). This view has however been counteracted by at least as many studies that claim health does indeed form a continuum (Heidrich, Liese, Löwel, & Keil, 2002; Manderbacka et al., 1998).

In responding to the 5-point Likert SRH question, the respondent is forced to combine evaluations on all the of the previously mentioned health aspects into a numerical value from 1 to 5, even when feeling healthy in one aspect, and less so in others (Jylhä, 2009). The self-rating of health is made possible by health knowledge and the corresponding illness theories or schemata, through which the individual translates bodily experience and sensory information into health-related categories (Bishop, 1991; Bjorner et al., 1996; Leventhal, Diefenbach, & Leventhal, 1992). In everyday life, abstract and theoretical medical information on disease may even be interpreted through somatic symptoms: the bodily experience guides the health views of the individual (Leventhal, Diefenbach et al. 1992). Jylhä (2009, p. 308 states: "Self-rated health differs from most indicators of health in that its origins lie in an active cognitive process that is not guided by formal, agreed rules or definitions." These processes may alter between intuitive and reflective.

4.2 Subpopulations and the processes of health self-ratings

As the previous chapter states, both health and health self-ratings are multi-dimensional. In addition to the individual details and circumstances of the respondent's health and personality, SRH depends on how the respondent interprets the study situation and the question (Krause & Jay, 1994; McDowell, 2006, p. 25). In survey studies some variation in reporting styles is considered expectable, if the variation is random. Following the *social comparison theory*, different social groups might have a systematically different understanding of either which aspects of health should be included in SRH, or different reporting standard or cut-off points. From here on in this study these reporting differences are called *reporting heterogeneity* (Dowd & Todd, 2011; Shmueli, 2003), though alternative terms have also been used.

The study of reporting style differences in health has been around for almost as long as the SRH measure itself. Zborowski (1952), followed by Zola (1966; 1973), alluded to cultural differences in symptom reporting. So far studies have suggested that at least age, sex, and cultural environment may affect health self-rating (Blaxter, 1990 13; Idler, 1993; Jylhä, Guralnik, Ferrucci, Jokela, & Heikkinen, 1998; Krause & Jay, 1994) as possibly does one's attitude towards one's own health (Idler et al., 1999). The first wave of this study tradition focused on sex and age or comparing nations; socioeconomic group differences, regions and ethnic groups followed soon after.

Reporting heterogeneity is most often explained with the element of social comparison included in self-rated health; both to the prevailing social norms and expectations of health, and to the observable health of others. Both may vary between population sub groups. According to Idler and colleagues (2004, p. 351), SRH reflects knowledge "--which is developed and refined by the first-hand experience of illness and honed in social settings where comparisons can be made." In Jylhä's (2009) conceptual model of self-rated health, individuals first consider the biomedical and functional health and health risks, then make comparisons with a chosen group, and finally add cultural and social norms and beliefs concerning health.

Social norms and the correlating health expectations (Baron-Epel et al., 2005; Delpierre et al., 2012; Singh-Manoux & Marmot, 2005; Zola, 1973) and expectations in the face of illness (Groot, 2000; Leinonen, Heikkinen, & Jylhä, 2001; Shmueli, 2003) may in part affect health ratings. What is interpreted as a normal state by some could to others be a symptom of a disease that requires intervention (Sen, 2002; Zola, 1973). Subjective health evaluations are made in the light of the best or the worst health states the individual can expect based on peer group observations and his or her own medical history and health experiences (Adams & White, 2006; Karisto, 1984, p. 23-24). Two people with the same level of for instance mild mobility limitations may be inclined to rate their status differently, if the other happens to be young and the other very old. For these reasons, several studies have questioned the comparability of the measure across different population sub groups, or between regions or countries (Groot, 2000; Huisman, van Lenthe, & Mackenbach, 2007; Lindeboom & Van Doorslaer, 2004).

Salomon and colleagues (Salomon, Tandon, & Murray, 2004) illustrate comparing SRH between populations: "*These responses are incomparable because the individuals have different response category cut points.*" D'Uva and colleagues (Bago d'Uva, Van Doorslaer, Lindeboom, & O'Donnell, 2008) continue: *"Homogeneous reporting behaviour corresponds to the assumption that the mapping [of the association between the* *latent and the measurable health variable] is constant across individuals. By contrast, reporting heterogeneity translates into different mappings between the latent variable and the observed categorical variable*."

Thomas and Frankenberg (2002b) present a three-dimensional model of health reporting. First, there is the individual factor: whether the person is more, or less likely than others to report a health condition or a problem. Secondly, there is the specific indicator of health: whether it is symptomatic or asymptomatic reflects on the likelihood of it being reported. Thirdly, there is the interaction component of the individual and the health indicator: for instance, health knowledge may vary between population subgroups and influence health-related reporting. The personal may either modify the connection between indicators of health and SRH or have an independent effect on the assessment. This modifying effect may be particularly strong in the case of diseases with little or vague symptoms.

For the young, symptoms and physical fitness, (Jylhä et al., 1986) and health behaviours (Fylkesnes & Førde, 1991; Krause & Jay, 1994; Manderbacka, 1998) have been found to be important factors for SRH, while for the middle-aged mental well-being and symptoms and for the oldest chronic diseases (Krause & Jay, 1994) were primary. Limitations in functional ability have also been suggested as important for the elderly (Moum, 1992). However, one study contradicts these findings (Borawski et al., 1996) stating that aging increases the likelihood of basing health assessments on health behaviours and health attitudes instead of symptoms or functioning. One study (Kaplan & Baron-Epel, 2003) found both age and the status of self-rated health itself to affect which dimension for judging health was chosen. The very old tended to compare their health to that of their peers, while the young only did so when they experienced their health as excellent.

The previously mentioned findings imply that the elderly and the chronically ill may choose different groups for comparisons than their young and healthy counterparts. Several quantitative studies conform (Cockerham, Sharp, & Wilcox, 1983; Ferraro, 1980; Idler, 1993; Maddox, 1962; Moum, 1992; Vuorisalmi et al., 2005) to these findings: the oldest of the old either report similar, or in many cases better SRH than their closest younger age group. This is often explained with the *reference group* theory. Self-rated health is dependent on the choice in comparison groups. It may be common for individuals to find a way to make more positive health evaluations by choosing the right peer group for the purpose. When ill, the elderly may choose to view their peers as the reference group (Levkoff, Cleary, & Wetle, 1987), a combination of peers and own health expectations (Cockerham et al., 1983) or personal health history (Suls, Marco, & Tobin, 1991). Idler (Idler, 1993) produces three explanations: SRH becomes more positive as the respondents age (the aging hypothesis), people born earlier are tougher (the cohort hypothesis) or simply that since SRH those with low SRH die, age eventually becomes positively associated with SRH (the selective survivor hypothesis).

In addition to age, reporting heterogeneity by sex has been studied in various contexts. Previous studies have found that after adjusting for medical indicators, any sex differences in health reporting styles tends to either disappear (Anson, Paran, Neumann, & Chernichovsky, 1993) or turn upside down from positive male reporting styles to females reporting better health (Ferraro, 1980; Idler, 1993). Some studies on the other hand have found sex to be the demographic factor that determined *which* biological indicators of health most affected subjective health evaluations (Goldman et al. 2004). This phenomenon has been explained with cultural gender expectations: women may be more likely to incorporate emotional health and a wide variety of other factors not directly relevant to life-threatening diseases into SRH (Benyamini, Leventhal, & Leventhal, 2000). Also, men may be expected to have a higher upper threshold for factors such as pain due to masculine socialization (Courtenay, 2000).

Still, when assessing results such as those by Goldman and colleagues (2004), one must bear in mind the prevalence of the various health indicators utilised in the study. Men are at a higher risk for cardiovascular diseases, and therefore this would be the disease group to affect men more than women. Hardly any studies have been able to account for genetics in their analyses. Future research should also account for the distribution of the both self-rated health and any other variables and covariates under scrutiny. How men and women interpret health may also affect their health behaviour and knowledge of their baseline health status. (Idler et al., 2000.)

Reporting heterogeneity may also be an issue when making comparisons between two cultures, for example two nations. While comparing SRH of the elderly residents of Tampere, Finland and St. Petersburg, Russia, Vuorisalmi and colleagues (Vuorisalmi, Pietilä, Pohjolainen, & Jylhä, 2008) found some evidence of cultural response styles. In both cities, SRH was associated with several indicators of health status, only positive health assessments were more common in Tampere even when selected symptoms, chronic diseases and indicators of functional ability were adjusted for. In other words, those in good health by the selected clinical measures were more likely to report good subjective health in Tampere.

Reporting heterogeneity has also been studied in relation to ethnicity. When comparing the Australian aboriginals with the non-aboriginal Australian population, the aboriginals tend to report better subjective general health than the general population, contradictory to the chosen objective health measures whereby the Aboriginals were the disadvantaged population subgroup (Mathers & Douglas, 1998). For Latinos living in the United States, SRH's predictive power for health outcomes was weaker for those who had lived in the country for a shorter period (proxy for acculturation) (Finch, Hummer, Reindl, & Vega, 2002). Still, similar results were not reached when studying the ethnic minorities of Great Britain. SRH was concluded to be valid across the sub-populations (Chandola & Jenkinson, 2000).

5. Previous studies on educational reporting heterogeneity

5.1 Self-rated health, education, and mortality

As stated in chapters 3 and 4, the elements survey respondents choose to include in their self-rated health assessments vary, and this variation may be systematic. Considering the arguments presented in chapter 4, reporting heterogeneity may be an issue when using SRH to compare the health status of different educational groups. Regardless, self-rated health is often used in the study of educational health disparities.

Several researchers have studied whether SRH is robust across educational groups by testing SRH's predictive power for mortality. By 2016, the association between self-rated health and mortality has been firmly established in over a hundred studies (DeSalvo, Bloser, Reynolds, He, & Muntner, 2006; Idler & Benyamini, 1997; Kaplan &

Camacho, 1983; Mossey & Shapiro, 1982). If death from a disease is the absolute worst state of health, comparing whether SRH predicts it equally across educational groups could partially answer the question concerning reporting heterogeneity (Ferraro & Kelley-Moore, 2001).

Whether the SRH-mortality -association is affected by the education of the respondent is still contested. To my knowledge, at least 3 studies have addressed the question, only to arrive at conflicting conclusions (Dowd & Zajacova, 2007; Huisman et al., 2007; Regidor, Guallar-Castillon, Gutierrez-Fisac, Banegas, & Rodriguez-Artalejo, 2010). Some studies have contributed to the discussion indirectly, as a by-product of studying something else (Jarczok et al., 2015). In addition to educational reporting heterogeneity per se, several studies also exist on the effects of socioeconomic positions on the association, as defined by for example the respondent's occupational class (Burström & Fredlund, 2001; Singh-Manoux et al., 2007) or income (Van Doorslaer & Gerdtham, 2003) instead of education (Subramanian & Ertel, 2008).

Most these studies are concerned with whether SRH is an equally valid tool for measuring health disparities as mortality. Huisman and colleagues (Huisman et al. 2007) concluded that it was likely that SRH worked equally well in the study of socioeconomic health differences as mortality, although it's predictive ability was found to be stronger for men in the highest educational group when compared to men in the lowest. For women, no such effect was found. Regidor and colleagues (Regidor et al., 2010) also found an interaction for men only.

Meanwhile, Dowd and Zajacova (2007), arrived at a similar conclusion using the NHANES data and dichotomised SAH -wording of self-rated health to examine variation in all-cause and cause-specific mortality. The researchers found that lower education as well as lower income levels hindered the predictive force of the SAH. Education was categorised as less than high-school, high-school and beyond high-school. Poor/fair health was a predictor for mortality in all educational groups, but the association was strongest in the highest attainment group. For example, the relative all-cause mortality risk was 2.82 (95% CI 2.67-2.98) in the education above high-school -

group, while it was 1.79 (95% CI 1.73-1,86) in the group that had not finished high-school. Similar patterns were found for cause-specific mortality.

So far, the results are mixed and the study settings varied. How education is categorised often varies. Also, some of the studies have used education as a solitary proxy for socioeconomic standing, while others have included occupation and/or income. SRH is worded in a multitude of ways. For example, Huisman and colleagues' (2007) SRH measure wording was unusual with two neutral options: "fair" and "sometimes good and sometimes poor" health. Regidor and colleagues (2010) used only the dichotomised results from the SAH wording, and included education as the single proxy for SES. None of the studies included education in their title.

When SRH is studied in relation to disease-specific mortality, relevant items such as comorbidities, disease severity, sociodemographic and psychosocial factors need to be adjusted for. Excluding relevant background variables may lead to an overestimation of the effects of SRH or SAH (Bosworth et al., 1999). Several studies presented here pointed out the need for the adjusting for comorbidities. For example, Huisman et.al. (Huisman et al., 2007) controlled for both pre-existing life-threatening and non-lethal disease and conditions, but utilizing data derived from self-reported conditions only. In the Dowd and Zajacova -study (Dowd & Zajacova, 2007), controlled factors were severe activity limitations, bed days during the past year, restricted activity during the past 2 weeks, self-reported BMI, age, sex, race, region, marital status, and household size. Apart from the BMI measure, the researchers did not disclose how the data on pre-existing conditions was gathered. Since the study data is derived from the National Health Interview Survey (NHIS) 1986–94 (Dawson & Adams, 1987), the measures are most likely self-reported. Adding activity limitations and BMI to the model lowered the effect of SAH on mortality, but did not remove it. Whether these are the most effective and valid controls can be contested.

According to Quesnel–Vallée (2007), the respondents in higher educational standing may incorporate more clinical, mortality-relevant information in addition to their subjective experiences to SRH. These differences may produce stronger association to mortality. Addedly, lower educational and socio-economic groups are at greater risk of death from external causes that are less associated with current health, and are also difficult to predict. Such causes of death are for instance accidents, and violence and alcohol-related incidents.

5.2 Diseases, biomarkers, functional health, and health behaviours

Mortality has been widely used as an indicator of health status in testing the predictive power of SRH. Disease-specific mortality is unquestionably an objective measurement of the health of an individual (Quesnel–Vallée, 2007). Nonetheless, as Jylhä (2009) points out, years lived healthy and free of disease are an equally important study topic. Mortality is also blind to the health-related quality of life and functional health –aspects included in the SRH (Moussavi et.al., 2007).

Examining the relationship of objective health indicators such as diagnosed medical conditions, biomarkers, biological factors, and health behaviours by education is not a novel idea as such. It is still, however, rare to focus on educational differences in SRH reporting styles, and even more so when using several health dimensions instead of focusing on just one, for example biomarkers. To my knowledge, six studies have addressed educational reporting heterogeneity and indicators other than mortality: five using SAH, the asymmetric global self-rated health wording (Delpierre, Lauwers-Cances, Datta, Lang, & Berkman, 2009; Dowd & Zajacova, 2010; Schnittker, 2009; Sohn, 2015) and two using SRH (Delpierre et al., 2012; Onadja, Bignami, Rossier, & Zunzunegui, 2013).

In the United States, Delpierre and colleagues (Delpierre et al., 2009) examined the association of poor health (dichotomised from SAH), with chosen physical self-reported conditions as confounded by education. Education was used as an indicator of socio-economic status. The self-reported conditions were experience of recent lengthened pain, duration of pain, presence of respiratory disease, a history of cancer or cardiovascular disease and functional limitations (assessed with a variety of scales, those reporting any difficulties were considered functionally limited) and trouble with vision, hearing or oral health. Models that accounted for health behaviours (smoking, alcohol consumption, physical activity, and BMI) were also constructed. Age, sex, ethnicity, marital status, and health insurance status were used as sociodemographic background variables.

Lower educational level was found associated with poor health, but when self-reported disease was controlled for, the lowest educational attainment group reported worse health. Simultaneously, SRH was associated with reported health conditions, but in the case of some of the health conditions, the association was stronger among the highly educated. An interaction effect was found for majority of the chosen variables. Only the associations between SRH and cancer and respiratory disease (both sexes) and chronic pain and hearing problems (males) were not modified by education. An interaction for functional limitations and oral health was found for both sexes and for respiratory- and cardiovascular disease and chronic pain for women only. The impact of education was therefore especially visible for women. The limitation of the study is, however, that all the health conditions are self-reported: the study compared self-reported single item health to specific self-reported health. The latter may already be influenced by education. The researchers themselves prompt adding objective health indicators to assessing the relation. (Delpierre et al., 2009.)

Delpierre and colleagues (Delpierre et al. 2012) continued on their previous work in France by adding psychosocial factors and household income to their work. In addition to the 5-point symmetric global SRH, subjective health was measured using the SF-36 health-related quality of life HRQoL -scale. Again, some self-reported health conditions, such as pains, had a stronger impact on SRH in higher education groups, while the relation was the opposite when the SRH measure was substituted by an indicator for general quality of life. However, the strength of the interaction depended on the indicator chosen to represent health. Again, Delpierre concluded that SRH may not be as stable as mortality in the study of health disparities.

Schnittker's (2009) main interest was in testing whether the association between reporting poor health (dichotomized SAH) and self-reported disability had changed over time in the US population above the age of 30. However, as a secondary question he analysed the association between poor SAH and 16 dichotomized self-reported functional limitations in daily tasks such as eating, bathing, dressing and walking. Education was found to strengthen all the associations, from the least strenuous tasks to the physically demanding ones. The strength of the interaction was notable. The study was based on the U.S. National Health Interview Survey (NHIS) data.

Onadja and colleagues (Onadja et al., 2013) utilised a three-dimensional health model. Conducted in Burkina Faso, the study compared dichotomised SRH to eight self-reported diseases, functional limitations as measured with the Short Set of Questions on Disability -survey tool, interviewer-diagnosed depression based on the depression module of the Mini International Neuropsychiatric Interview (MINI) and four health-related factors: BMI, current smoking, alcohol consumption exceeding two times per week and physical activity. The latter four were not articulated to represent health behaviours but are widely used as indicators of health behaviours in other studies (Manderbacka et al., 1999; Sohn, 2015).

The self-reported diseases were hypertension, diabetes, chronic bronchitis or asthma, angina pectoris, stroke, arthritis, gout, and stomach ulcer. Functional limitations (no/one difficulty/two or more difficulties) were inquired using four-point scales that measured the amount of difficulty the respondent was experiencing in relation to six different aspects of functioning, such as vision and mobility. The depression interview consisted of two questions concerning the respondents' overall mood and interest in activities, and further seven questions for any respondents experiencing a decline in either one of the previously mentioned. Reporting five or more symptoms resulted in a depression diagnosis. (Onadja et al., 2013.)

All the selected indicators were found associated with SRH, but the association between depression and SRH disappeared in the full model that included all the health indicators. The association was strongest for functional limitations. In the second stage of the analysis, the data was stratified by age, sex and then education (primary school/higher than primary school). No moderating effect was found for depression or self-reported disease, but functional limitations were associated with SRH in the group with the lesser educational attainments only. When disease was excluded from the model, functional limitations were found associated with SRH in both educational groups. The researchers concluded that disease is a confounder between SRH and functional limitations for educated residents of Ouagadougou. These results conflict with Delpierre’s earlier findings. (Delpierre et al., 2009; Onadja et al., 2013.)

Dowd and Zajacova (2010) addressed the question through SAH and biomarker data alone. They stated that biomarkers - unlike self-reported diagnosed illnesses which are at least partially dependent on the study participant’s habits of attending check-ups –are virtually free of systematic reporting error when measured in a laboratory for a random sample. The researchers tested 14 selected biomarkers and found a significant interaction effect between SAH and education in relation to 12 of these. From these findings, they drew that the SAH-objective health indicator -relation was modified by education. Those reporting good health only shared a biological risk profile if they also shared the same level of education as well. When biological risk factors and health indicators are taken into consideration, those with lower educational attainments and good self-rated health level fared worse than those with a higher education at the equivalent level of SRH. Education again functioned as a proxy for SES.

Dowd and Zajacova’s argument for favouring biomarkers to self-reported diagnosed illnesses is, indeed, appealing. However, Sohn (2015) notes that some of these biomarkers are related to physical symptoms while others are not: some are therefore more likely than others to reflect on the individual’s bodily experiences. Knowledge of certain biomarkers, for instance cholesterol, can be obtained by having a medical check-up. At least in the USA education, is associated with access to medical care. Individuals with higher education would thus be more likely to be aware of their health biomarker status. In addition, education can affect both the intensity and the swiftness of internalizing new medical information (De Walque 2004).

To my knowledge, Sohn (2015) is the first to analyse the modifying effects of education on SAH in relation to a wider concept of health - biomarkers, medical conditions, and health behaviours - by education. The relationship between subjective health and similar multidimensional health definitions has, however, been analysed before. For instance, Manderbacka (1998) and Jylhä (2009) based their articles on similar constructs, but did not include education as a modifying factor.

Sohn's data on medical conditions is based on self-reports of having received a diagnosis for any of the following: heart failure, diabetes, chronic obstructive pulmonary disease, liver condition, heart disease, heart attack, angina, and stroke. Participants were also asked whether they take medication for hypertension or high blood pressure. The selected health behaviours were current smoking, a history of heavy binge drinking and recreational physical activity. The biomarker data was derived from a clinical examination and consisted of HDL cholesterol, haemoglobin, C-reactive protein, resting heart rate, body mass index, waist-to-height ratio, and Hepatitis B measures. (Sohn, 2015.)

The study concluded that the relationship between the selected medical conditions and SRH did not disappear across educational levels. However, in the case of health behaviour and biomarkers, the strength of this association did vary: it was stronger for the most educated group. These findings are partially in line with Delpierre's previous work (2009; 2012).

Lindeboom and colleagues (Lindeboom & Van Doorslaer, 2004) addressed this matter by comparing SRH by population sub group to a wholly subjective health scale. The previously cited article did find some proof of heterogeneous response style by age and sex, but not by any other factors.

5.3 Summary and gaps in the evidence

Self-rated health is often thought to reflect some true, underlying state of health (Thomas & Frankenberg, 2002b). And yet, when Manderbacka (1998b) requested her interviewees to describe their health, they produced concrete, contextual, and partly contradictory responses. Despise these challenges, research does suggest (Idler & Benyamini 1997) that survey respondents still integrate a large amount of health-relevant information into SRH.

Per Thomas and Frankenberg (2002), there is always a degree of deviation between health self-ratings and "true" health, but the amount of the deviation may indeed be dependent on the factors presented in the previous chapters. In fact, there is a growing number of studies that address self-rated health and the modifying effects of the cultural context; a smaller number has addressed education. So far, the studies have not agreed on whether reporting heterogeneity is a problem when using SRH in the study of health disparities.

There are various limitations to the previously presented studies. Firstly, comparability is an issue. The exact wordings and option choices for SRH vary, as do the included correlates (Bjorner 1996, p. 60). It is not yet clear if the results obtained with self-rated health on a scale of excellent, great, good, fair and poor can or should be compared to those obtained using 5-point Likert scales where the neutrally expressed value of “fair” or “average” is the middle one (Lindeboom & Van Doorslaer, 2004). Several studies have also used dichotomised SRH or SAH. All and all, several researchers and social commentators have expressed concern over the possible differences in correlational structure of SRH, while few studies even partially back up these claim (for the discussion, see Chapters 3 and 4). The place of education in the studies is also an issue. The most common choice in previous studies has been to use it as a single designator for SES (Adams and White 2006).

Secondly, most of these studies have not addressed multi-dimensional health, but for example selected biomarkers or diseases only.

Thirdly, the two studies that come closest to the study questions presented here both utilise the same data, the American NHANES (Dowd & Zajacova, 2010; Sohn, 2015). Studies on effects of education on the association between SRH and multidimensional health have, to my knowledge, not been conducted in Scandinavia or Norther Europe, though researchers have suggested the topic for future study (Dowd & Zajacova, 2010; Jylhä et al., 1998; Jylhä, 2009; Sohn, 2015).

Fourthly, in the studies cited previously in this thesis, the study population sizes have varied notably from Maddox’s and Douglas’ (1973) compact 83 participants to Fylkesnes’ and Førde’s (1991) 18 560. In the four studies that come closest to my study questions, the sample sizes varied between Sohn’s (Sohn, 2015) 30,823 and Onadja and companion’s 2195 (Onadja et al., 2013). The sample size affects the statistical power of the analyses.

Age, health behaviours, self-reported chronic conditions and diagnosed clinical diseases all are widely accepted to affect health self-ratings. However, as the foregoing literature review suggests, it is yet not clear if education modifies these associations. Dowd and Zajacova (2010) are the strongest advocators for education as a modifier, at least in the case of the United States. Sohn (2015) takes the middle ground and presents that education is a modifier, but for health behaviours and biomarkers only.

6. Aims of the study

6.1 Research questions

The purpose of this Master's Thesis is to contributed to the discussion the basis of self-rated health, reporting heterogeneity and the study of educational health disparities.

This study will answer the following questions:

1. What is the relationship between poor or very poor self-rated health and selected measures and indicators of health?
2. Are these relations modified by educational status?

The selected health indicators represent the five health dimensions -concept presented in the previous chapters. The so-called negative definition of health as absence or presence of disease is predominant in the study: when used alongside data on subjective experiences and the functional dimension of health, this frame is practical, inclusive, and comprehensible (Karisto, 1984, p. 61).

6.2 Data

The data used in the present study is from the Health 2000, a nationally representative interview and health examination survey conducted in 2000-2001. The main purpose of the study was to produce a comprehensive picture of the Finnish population health. The target population consisted of household population and individuals living in institutions in mainland Finland, aged 18 and over. The data was comparable to the Mini-Suomi study conducted between the years 1978 and 1980 (Aromaa et al., 1989; Aromaa & Koskinen, 2004). The principal responsibility for Health 2000 was assigned to the National Public Health Institute (KTL), which later, on became the National Institute for Health and Welfare (THL).

The cross-sectional study data is multidimensional and includes both subjective and objective health measures. The data were collected by trained personnel consisting of interviewers, nurses, and doctors. First, Statistics Finland conducted a home interview. A date for a medical examination was settled at the time of the interview and the study participant was provided with an additional home survey questionnaire (questionnaire 1). The medical examination (including a symptom interview, a mental and oral health examination and laboratory and functional capacity testing) was conducted within a few weeks following the interview. The participants also filled an infection survey questionnaire while visiting the clinician and were provided with a second home survey questionnaire and a separate nutrition and diet questionnaire. For a more detailed description, see Heistaro (Heistaro, 2008).

Supplementary data collection methods were used to guarantee maximum participation rates. For those participants who could not attend the medical health examination, a succinct medical examination conducted at home was offered as an alternative. Substitutive shorter forms were made available for the full home interview and the basic survey questionnaire (questionnaire 1). The above-mentioned home survey and medical examination formed the main body of the study. Those participants absent from these two sections were approached via telephone and were also sent a second survey form to gather basic information and some of the more crucial data. (Koskinen et al., 2008, p. 132-133.)

The survey was linked with additional data from various register sources. The subjects were informed of the linking in writing and signed a consent form. The linked data concerns such matters as income, causes of death, purchased and prescribed medicines and employment. Current regulations were taken into consideration in all record linkages and the linkage process was executed in cooperation with all the involved parties and following the directions and ethical regulations issued by the data protection authorities. The data was anonymised before I was granted access. (Koskinen et al., 2008, p. 180; Laiho, Djerf, & Lehtonen, 2008, p. 13-15.)

6.3 Sampling design

The population-wide insurance database was used as the sampling frame. A stratified two stage cluster design was utilized to draw a sample to accurately reflect the demographic distributions. Since the study consisted of an interview and a clinical examination, a complex sampling design was unavoidable to maximize the ease of attending the medical examination for the participants with reasonable expenditures. The overall participation rate for the study has been considered good. In the case of the aged 30 and over -study group, 89% of those sampled participated in the home interview section and nearly 80% in the general health examination. (Koskinen et al., 2008, p. 180; Laiho et al., 2008, p. 13-15.)

University hospital districts of Helsinki, Turku, Tampere, Kuopio, and Oulu were used for the geographical stratification. In the first sampling stage (clustering), 80 health centre districts were selected. The strata were divided into sub-strata so that the 15 largest cities in Finland were guaranteed selection. The respective sample size was defined as proportional to the population base. The remaining health centres were selected utilising systematic probabilities proportional to size (PPSY-SYS). The actual individual units, the population aged 18 or over, were sorted by age. The final stage of the selection was conducted using systematic random sampling that, due to the sorting, involved implicit stratification by age. (Laiho et al., 2008, p. 14-15.)

The sampling design for the main survey included oversampling (within clusters) with a double sampling fraction for people aged 80 or over to insure an adequate amount of data for the elderly. The original overall sample size was set at approximately 8000 and resulted in a sample size of 8028 (Laiho et al., 2008, p. 13-15).

The oversampling of the elderly (80 years and older) was corrected using sampling weights. Post-stratification weights were used to account for non-response. The least likely population groups to agree to take part in the survey were the youngest men, the oldest women, those with least education and income and those living alone. (Laiho et al. 2004; Djerf et al. 2008, p. 182-200).

6.4 Final sample and missing data

Table 1 presents the study participation for each separate stage of the survey. The data utilised in this study is combined from several of the sections: the interview, the home questionnaire, and the clinical health examination for those aged 30 and over, plus additional register data.

Table 10.1. Original sample, final sample, and participation in different stages of the collection of the data utilised in this study.

	n	%
Sample	8028	100
Deceased before fieldwork	49	0.6
Final Health 2000 sample	7979	100
Participation in one or more of the data collection stages	7415	92.9
Non-participation: refusal, abroad, other	564	7.1
Participants in some stage of the health examination	6354	79.6
Participants in home-visit interview	7087	88.8
Laboratory tests conducted	6354	79.6
Functional capacity measurements conducted	6329	79.3
Clinical examination finished	6326	79.3
Participation in mental health interview	6005	75.3
Participation in one or more of the data collection stages	7415	100
Completed self-assessed health	7364	99.3
Full demographic information	7332	98.9
Completed survey items	6563	88.5
Complete clinical and laboratory work	5967	74.3
Completed evaluation of subjective abilities, tests of abilities	5899	73.5
Completed mental health interview	5586	69.6

Out of the sampled 8028, 7415 participated in at least one stage of the study. 664 respondents did not reply to the question concerning self-rated health. Out of these 664, the clear majority (around 630 people) had agreed to participate but had either provided very few answers or none, whereupon all or most of their data was derived from registers. To be able to make comparisons, those respondents that had not attended one of the clinical examination options (health centre, home) and had not provided responses to all the survey items utilised in this study were excluded. This resulted in the final sample size of 5586.

All though the quality of the Health 2000 data is generally considered to be good, the participation rates in the original sample varied somewhat between different socio-demographic and socioeconomic groups. Even after the telephone-assisted interviews and the secondary questionnaire, the long-term unemployed, those with a small income, non-Finnish speakers, and those with an unclassified family- and socioeconomic status were least likely to participate (Koskinen et al. 2008, p. 132-143).

After the selections made for this study, missing data analysis was conducted using the Stata *misstable* -set of procedures and logistic regression analysis. For 72 % of the full sample of 8028, none of the variables had missing data. The second most common pattern for missing values was all values missing, which consisted 8 % of the original sample of 8028. After these two larger groups, 3 % and 3 % and 2 % of the 8082 had between 7 and 3 values missing; mostly in the section of the subjective tests. The cut-off point requirement of participating in all the stages of the study was found to be less likely for those reporting poor or very poor heath (OR .45). Education was not found associated with the likelihood of having participated in all the stages, but higher age and male sex decreased the likelihood of participation in all stages.

6.5 Outcome variables

The outcome variable *self-rated health* (item BA01) was included in the home interview section of the Health 2000 -study. The question was formulated as follows: “Would you describe your current state of health as -” and the answer options: “Good”, “fairly good”, “average”, “fairly bad” and “bad”, with 1 representing the best state of health and 5 the worst. A dichotomous variable was formed with options 4 and 5 indicating poor health and 1-3 better than poor. Similar dichotomy has been utilised in several studies (Bjorner et al., 1996), while sometimes “average” health is grouped together with poor and very poor health. In the preliminary analyses the variable was also used as is.

6.6 Health measures

The section of the study from which the indicator data has been derived from is presented in table 1 for each chosen variable. The frequencies for the measures and indicators are displayed more thoroughly in Appendix 3. The data utilised here is derived from the home interview (both long and short), home questionnaire, medical examination and register data. The self-rated health item and the smoking items were included in both the long and short home interviews and therefore both were admitted.

The health indicators have been derived from several sections of the study (for instance the survey or the interview, see: 5.2 Data and Table 1). Information concerning smoking and vegetable consumption habits was collected by interview, while exercise and alcohol consumption were addressed in the home survey questionnaire. The indicators addressing clinical health (specific disease diagnosis) were formed based on analyses conducted centralized at the National Public Health Institute (KTL) department of health promotion and prevention of chronic disease and illnesses (ETEO). For more information and the original study forms, please visit: http://www.terveys2000.fi/data.html.

The medical examination refers to both the main clinical examination and the alternative home visit by a medical professional. In total 408 participants chose the option of a home visit by a clinician instead of attending the medical examination. During the home visit, all other clinical measures addressed in this study were taken into consideration except for evaluations for musculoskeletal conditions and clinician’s overall evaluation of functional abilities. Objective functional ability tests such as hand grip strength were conducted.

Table 1. Study data: stages of the Health 2000 research process.

	Data source				
Health measure	Home inter-view, full/short	1st home questionnaire	Medical ex-amination	Register data	Other sections
SRH	x				
Smoking	x				
Alcohol consumption		x			
Physical activity		x			
Diet		x			
Cardiovascular	x		x	x	
Hypertension	x		x	x	
Diabetes	x		x	x	
Musculoskeletal dis-orders	x		x		
BMI		x	x		
Cholesterol			x		
Objective functioning			x		
Subjective functioning	x				
Mental health			x		
Somatic complaints		x			
Sex				x	
Age				x	
Education				x	
Income	x				
Employment status	x				
Marital status	x				
Residency				x	

The health variables used as indicators of health in this study have been chosen on the basis of both prior research (Blaxter, 2003; Idler et al., 1999; Kaplan & Baron-Epel, 2003; Thomas & Frankenberg, 2002a) and the possibilities and limitations of the data. The selected variables have been divided into five categories, representing different dimensions of health: health behaviours (health as doing), objective clinical health and biomarkers (biomedical health), functional capacity (functional health), mental health and bodily sensations and pain.

6.6.1. Health behaviours

Five measures were selected to represent health behaviour on the basis of previous studies: smoking, alcohol consumption, diet and leisure time physical activity (Blaxter, 2003, p. 5; Heistaro et al., 2001; Manderbacka et al., 1999).

Smoking. In the Health 2000 study smoking status was measured using 11 questions on smoking habits, for example items distinguishing regular smokers from occasional- and non-smokers and items inquiring the volume of tobacco consumption. In this study, dichotomised variable for daily smoking (smokes daily/does not smoke daily) is used, formed based on the replies on all the tobacco variables. This choice follows the EHRM recommendations for primary indicators for smoking in survey studies (Tolonen, Kuulasmaa, Laatikainen, & Wolf, 2002).

Leisure-time physical activity. The data was derived from two survey questionnaire items measuring the respondent's physical activity: how many times per week he/she engages in any leisure time physical activity that is intensive enough to cause at least mild shortness of breath for the minimum duration of 30 minutes and whether he/she bicycles or walks to work or back. The survey items did not specifically instruct to include other information than this, for instance physical labour or performing intense chores. On this basis, the respondents were categorised into four groups by the Health 2000 staff following the current recommendations on physical activity of any kind for 30 minutes 4 times per week. These groups were: ideal-, sufficient-, indefinite- and inadequate activity level. These categories were further dichotomised into inadequate activity/not inadequate (ideal, sufficient, indefinite).

Dietary habits. Eating habits were inquired in Health 2000 with several questions addressing for instance vegetable and sugar consumption. Vegetable consumption is a widely-used proxy for diet. Consuming fresh vegetables decreases the risk of cardiovascular diseases, some cancers and adult-type diabetes. The question was posed as follows: 'How often have you eaten vegetables during the last week?' The alternative answers were 'never', 'once or twice', '3−5 times' and '6−7 times'. The responses were dichotomised following the National Institute of Health Welfare Compass cut-off points into those who ate fresh vegetables two days per week or less often and those exceeding this frequency (Terveyden ja hyvinvoinnin laitos, 2016).

6.6.2. Clinical health

The second health category is clinical health: observable clinical conditions and biomarkers. The existence of a disease was identified by the Health 2000 team using several sources: the medical health examination and with additional data from the home interview, the symptom interview, and relevant registers such as the HILMO-system and medication registers from 1964 to 2004. Based on the previously mentioned, disease status was categorised by the study team as: "no diagnosis", "possible diagnosis", "certain diagnosis" and "patient has received medication/diagnosis for disease after 2000/2001 as registered in the HILMO-system". The final dichotomized disease items used in this study were formed from this basis. More information on the diagnoses and survey item wordings can be found in Appendix 3.

Heart infarct, *heart failure* and *stroke* were combined into a single item labelled *cardiovascular conditions*. In Finland, mortality due to cardiovascular diseases has decreased since the 1970's, but still, cardiovascular disease remains among the most common causes of death, while regional and socioeconomic differences in occurrences are notable. Risk factors include hereditary factors, health behaviours (diet fat content and smoking), hypertension, diabetes, and high cholesterol levels; also, the risks increase with age and are higher for men. (Reunanen, 2005; Tarkiainen, Martikainen, Laaksonen, & Valkonen, 2012.)

Heart failure is not a disease, but a symptom related to several heart diseases, often coronary artery disease and infarct. The heart still functions but at a lessened capacity so that the body's oxygen needs are no longer fully met. The symptoms vary by which chamber of the heart is afflicted and may manifest in shortness of breath, fatigue, dry cough, and liquid build up in tissue. Left-sided heart failure may acutely result in pulmonary oedema, which requires immediate medical attention. Heart failure may also increase the risk for stroke. (Lommi, 2013.)

Heart infarction is usually the result of a long-standing coronary artery disease which has narrowed the arteries, finally resulting in a block which reduces or completely cuts off the blood flow, causing damage to the heart muscle affected by the artery. It is a severe condition, possibly fatal. (Kervinen, 2013.)

Strokes are caused by blockage (ischemic stroke, more common) or bursting (haemorrhagic stroke) of a blood vessel, so that the blood flow to some area of the brain is prevented. The oxygen deprived area then begins to die (infarction). The results vary upon where the damage occurred and the extent of the damage: from loss of control in limbs to losses in functional and cognitive abilities and which may or may not be permanent. Stroke requires immediate medical attention. (Roine, 2013.)

Diabetes (diabetes mellitus) is characterised by chronic hyperglycaemia along with metabolic disturbances caused by either insulin secretion, insulin action, or both. Diabetes has both immediate and chronic health consequences. The daily symptoms may vary from nearly asymptomatic to such problems as weight loss, fatigue, thirst and blurring of vision. In its most severe forms and if untreated a life-threatening ketoacidosis or a nonketotic hyperosmolar state may develop. Eventually, diabetes may advance dysfunction or even failure of several organs and increase the risk for processes that lead to neuropathy (often affecting the lower limbs), retinopathy and blindness, ulcers and cardio-, peripheral-, and cerebrovascular disease. Diabetes is often classified into type I and type II; the latter is slow to develop and is primarily treated with lifestyle changes. Both types are associated with all the above mentioned associative diseases and dysfunctions. In this study the existence of both types I and type II diabetes was confirmed in the clinician's evaluation and on this basis a dichotomised variable was formed: the patient either does or does not have either type. (Alberti & Zimmet, 1998; Yki-Järvinen & Tuomi, 2013.)

Data concerning the existence of *musculoskeletal disorders* was also derived from the clinician's evaluation. All possible disorders (neck, back) were recoded into a new dichotomous variable. Out of the diagnosed disorders, 250 were related to neck (vertebrae fracture and whiplash related issues, cervical syndrome, back rheumatism, etc.)

and 373 to lower back (vertebrae fracture, spinal stenosis, sciatica, NUD, etc.) diagnoses. Unclear cases (n=34) were recoded as negative diagnosis.

Several *biomarkers* were selected to represent risk profiles and pre-clinical health: HDL and LDL cholesterol, blood pressure and body mass index (BMI). A trained medical staff collected all the measurements. The biomarker data was again dichotomized, following current guidelines.

A respondent was identified with *high cholesterol levels* when his/her total cholesterol level exceeded the European guideline threshold level for very high cholesterol (total above 6.2 mmol/L) and either his/her HDL cholesterol or LDL cholesterol was simultaneously categorised as highly undesirable (below 1 mmol/L or above 4,9 mmol/L, respectively). Recommended optimal maximum cholesterol was tried as an alternative cut-off point. Since almost three out of four of the participants (and working age Finns) exceed these recommendations, the very high –levels were used instead of the recommended lower limit. (Reunanen, 2005.)

High blood pressure was identified by combining data on blood pressure measures and self-reported blood pressure –related medication and register data for special reimbursements gathered during the clinical examination. A combination of either one of the previous two and blood pressure exceeding current medical guidelines was interpreted as a positive outcome. Since to reflect a persistent condition or hypertension blood pressure needs to be elevated for a sustained period and therefore single survey blood pressure measurements alone are not recommended for screening (Tolonen et al., 2002). It is notable that - like with cholesterol recommendations - since current medical guidelines of systolic blood pressure average below 120 mmHg and diastolic blood pressure average below 80 mmHg were used as cutting-off points for the binary variable, almost half of the study participants were diagnosed with hypertension (Piukka, 2008).

Body height and weight were measured during the health examination section of the study and used to calculate *body mass index* BMI. In case the health examination information was missing, BMI was calculated using self-reported height and weight (home survey section). The thresholds for normal BMI are the same for both sexes, although body proportions are relevant for the interpretation. In this study, BMI was dichotomised using the WHO (World Health Organization, 2016) and EHRM (Tolonen et al., 2002) recommendations for principal cut-off points for obesity at BMI>=30 (obesity classes I-III).

6.6.3. Functional capacity

The fifth health dimension is functioning. Both objective and subjective functional health indicators were addressed.

Subjective functional capacity concerns the level on which the respondents claim their social and role functioning to be limited by their mobility, functional abilities and/or health. In the Health 2000 these measures were collected in the home interview section of the study (E-K). Most of the single items that address functional health and capacities are derived from disability and ability scales of a scale type often referred to as Activities of Daily Living (ADL) -scales and later extensions known as Instrumental Activities of Daily Living (IADL) or Performance Activities of Daily Living (PADL). Early versions of the ADL-scale have been utilised since the 1950's and typically included the subject's level of independence in such basic functions as eating and dressing, but during the late 1960's the concept was extended to include items addressing applied problems such as difficulties with shopping or managing ones' finances. Conceptually, these latter constructs are close to the idea of health as the ability to fulfil social role functions, and suitable when targeting the general population instead of just the elderly or the severely disabled (McDowell, 2006, 56). In addition to daily living, several items in the survey addressed subjective physical condition and fitness in performing tasks such as running and climbing stairs.

The Finnish National Institute for Health and Welfare's Health Compass (Terveyden ja hyvinvoinnin laitos, 2016) recommends the following measures for subjective functional ability: for subjects aged between 20 and 64 the ability to run 100 meters, for those aged 65+ walking 500 meters and for the age group 75 and older conducting daily affairs. The Health 2000 data did not include an item on running, so the variable was replaced with the ability to climb several sets of stairs without rest (cardiovascular effort). Separate measures are used for the age groups so that the measure would reflect limitations typical for each group. Young respondents with difficulties with walking should be very rare, while troubles with climbing stairs or running could already occur due to the weakened state of either the musculoskeletal-, the respiratory- or the vascular system. Early decline in fitness and mobility may prognosticate future limitations in mobility. For the elderly, having trouble with stairs should be rather common. The ability to walk 500 meters, for instance to buy groceries, is vital for unassisted living as is the ability to manage in daily tasks. The selected items addressing functional capacity was also dichotomized (able to perform age task/unable or find it somewhat or very difficult).

The questions were posed using capacity wording: does the respondent believe he/she would be able to perform the task should he/she choose to try it. The respondent is not asked about performing the task regularly. The capacity wording is considered positive and less conservative in relation to lifestyle choices. It measures health impairments instead of handicaps. (McDowell, 2006, p. 15-16.)

Objective functional capacity. For the participants aged 30 to 54, handgrip strength (HSG) was selected to represent objective functional capacity. Out of a variety indicators that represent muscle strength, handgrip strength is simple and its use is widespread. It is a recommended general indicator of strength (Bohannon, 2001); strength in turn is important both for its relation with functional capabilities and future health outcomes. It has been confirmed a predictor for mortality (Cooper, Kuh, & Hardy, 2010; Rijk, Roos, Deckx, van, & Buntinx, 2015), future disability (Bohannon, 2001; Rantanen et al., 1999; Xue, Walston, Fried, & Beamer, 2011) and future functional and mobility limitations (Rijk et.al. 2016) at least with the elderly populations.

Handgrip strength was measured using a handheld Good Strength IGS01 dynamometer measuring device provided by Metitur. Following the current recommendations, the participant was instructed to grip the device handle as hard as possible for the duration of 3-5 seconds, using the dominant hand. The test was repeated after 30 seconds to allow muscle recovery; if the difference between the two tests exceeded 10 %

the test was repeated one more time. The maximum performance was recorded in newtons. Sex, age, and body composition affect the results. Several studies have suggested cut-off values for analysing the indicator. Here the test results were dichotomised per the age and sex-specific cut-off points recommended by The National Institute of Health and Welfare's guidelines into two groups: results severely below group averages and results below, equal to or above group averages. (Rijk et al., 2015; Sainio, Aromaa, & Koskinen, 2008.; Viitasalo, Era, Leskinen, & Heikkinen, 1985.)

Different tests for functional capacity were conducted for participants over the age of 55 as recommended. Maximal walking speed test was the selected test measure for this age group. The test times how many seconds it takes the subject to walk the distance of 6.1 meters. The results were dichotomised following the cut-off points recommended in The National Institute of Health and Welfare's guidelines into two groups: results severely below age-specific averages and results below, equal to or above age-specific averages. (Sainio et al., 2008.)

The final functional health measure is the *clinician's evaluation of the subject's functional abilities* after the medical examination and the clinician's interview. It accounts for not just the above-mentioned abilities, but also hearing, eyesight, basic cognitive abilities and speech and is evaluated in relation to the patient's age group. The functional abilities were classified into four groups: not diminished, slightly diminished or sensory limitations, substantially diminished or several sensory limitations and fully or almost fully diminished functional abilities. The results were the dichotomised as diminished abilities yes/no, with slightly, substantially, and severely diminished abilities as a positive outcome and no limitations as a negative.

6.6.4. Mental health and psychological well-being

The final category is psychological and psychosocial well-being. The health indicators selected for this category are mental health disorders as diagnosed in a clinical interview setting using the Composite International Diagnostic Interview (CIDI) tool for screening for DSM-IV -classified psychiatric disorders in the general population.

The CIDI is developed and maintained by the World Health Organization. It is a face-to-face survey interview tool that was in the case of Health 2000 executed as a separate section of the clinical examination. The CIDI collects information about mental disorders, psychopathologies, alcohol and substance abuse and their treatment (or unmet needs thereof). Because of the rarity of individual mental conditions that the CIDI instrument can be used to recognize, the multiple CIDI diagnoses (panic attacks, social phobias, agoraphobia, anxiety, depression, dysthymia, alcoholism, or comorbid combinations of any of the previously mentioned) were dichotomised into diagnosed mental disorder/no diagnosed mental disorder. (Kessler & Üstün, 2004.)

6.6.5. Bodily sensations and pain

This health dimension consists of bodily experiences such as pain and ache and has been suggested as an important factor in relation to SRH since it pictures sensory information that is only available to the respondent (Jylhä, 2009). The frequency and intensity of somatic complaints were measured using the SCL-90 tool. The full tool is an inventory of 90, mostly psychological symptoms which are divided into nine dimensions of which the first is somatisation. This section of the tool originally comprises of 12 questions concerning complaints such as headache, nausea, and various pains which the respondent rates on a scale of distress from 0-4. The Health 2000 survey also included a question on general pains and aches but it was omitted in this study and the tool was used in its original form. The items were all formulated as follows:" Next we will address your recent symptoms. To what extent are you troubled by --", the answer options available being: “Not at all, very little, somewhat, quite a lot and very much”. Raw somatisation scores were calculated by dividing the sum of scores by the number of the items the respondent had completed and from there a t-score can be calculated. Per recommendations, should t-score be equal to or greater than 63 or the participant belong to the 90th percentile within their norm group (sex/age/patient or non-patient population), the results warrant clinical inspection. The percentile -criteria was used here. (Derogatis & Unger, 2010.)

The tool is a subjective measure and due to the nature of the symptoms it covers, it may reflect for instance cardiovascular, gastrointestinal, or respiratory conditions, pre-clinical states or on the other hand a variety of anxiety disorders. While its stand-alone use has been contested in some studies, most still recommend it as a good measure of overall somatic distress (Woolfolk & Allen, 2007, p. 77-80). A recent systematic review rated it recommendable for large scale surveys (Zijlema et al., 2013).

6.7 Statistical analysis

All analyses and estimations were conducted using STATA 11.2 software. The research objectives were approached using descriptive techniques, regression analysis and regression -based decomposition methods.

Before fitting the models, basic descriptive sociodemographic statistics and the prevalence and distribution of the health indicators and coefficients were presented. The correlations between the health indicators and the outcome (poor SRH) were also assessed (for details see Appendix 1). Multicollinearity of the explanatory was tested by comparing the correlations (Appendix 1) and computing variance inflation factor (VIF) –values using the lowest possible conventional level for acceptability and keeping in mind the possible effects of large standard errors in the data (O'Brien, 2007). Any offending variables would have been removed before the analysis had any been discovered. None of the predictors were highly correlated with each other, although a moderate correlation existed between clinician's estimations of functional abilities, musculoskeletal and cardiovascular troubles, and subjective evaluations of functional troubles in every-day activities (0.0-0.41, the highest correlation being between clinician's estimations of functional limitations and musculoskeletal troubles). The correlation and low VIF-values were interpreted as acceptable. (Gujarati, 2003, p. 596-597.)

Age-adjusted prevalence rates of the dichotomised self-rated health and other variables were estimated for each educational and self-rated health group using predicted margins (Graubard & Korn, 1999). The sampling design is complex and thus all analyses need to account for the stratification and clustering as well as the sampling weights. The National Institute of Health has published recommendations for the data (Djerf, Laiho, Lehtonen, Härkänen, & Knekt, 2008) so that these complexities can to a reasonable extent be accounted for. Also, all descriptive estimations were adjusted for age and run for men and women, both separately and together, although here only the pooled results for both sexes were presented.

The first part of the analysis was conducted using logistic regression, a transformation based on linear regression, and the second part using logistic regression based nonlinear decomposition. Logistic regression models the association between a binary dependent variable and possibly multiple explanatory variables (either categorical or continuous) by maximum likelihood, in other words by evaluating the probability of a positive outcome of the dependent variable. Logistic regression is suited to the study question, because it allows for models with several contemporaneous explanatory variables. Therefore, the question of which of the variables is associated with the dependent variable when other known explanatory factors are present and the strength of this association can be addressed. (Garson, 2012; Gujarati, 2003, p. 596-597.)

In the first part of the analysis, three logistic regression models were fitted. In the baseline model each health indicators though to affect health self-ratings was included separately, while age and sex were controlled for. In model 2, all the explanatory health variables were addressed simultaneously while controlling for age and sex. In model 3, education was added to the previous. Results are presented in Chapter 7.

The formula for the logistic regression can be simply presented as such:

(1.1) $$\ln\left[\frac{P(Y=1)}{1-P(Y=1)}\right] = a + bx$$

, where P(Y=1) denotes the probability of the binary outcome variable being 1. The equation for the first variable of the Model 1 could be written as:

(1.2) $$\text{logit}(P_{SRH}) = \alpha + \beta AGE + \beta BGENDER + \beta SMOKER + \varepsilon$$

, where when the outcome variable SRH receives the value 1 if the respondent reports poor health.

Stata fits the logistic regression model using the following (Stata Manual, 2016):

$$(1.3) \qquad \Pr(y \neq 0 \mid x_J) = \frac{\exp(x_J\beta)}{1 + \exp(x_J\beta)}$$

From this, an antilog is taken to calculate odds ratios and the equivalent standard error ratios and from there the confidence intervals. The results are presented as odds ratios (OR), which are computed by calculating the odds of the positive outcome (poor health) for individuals in both the presence and the absence of the health indicator variable and then dividing the former with the latter to obtain the ratio. The interpretation is therefore to compare the relative odds of the occurrence: the odds when the indicator is present to the baseline of it being absent. In other words, ORs were selected to facilitate interpretations. Results are presented with the conventional 95% confidence intervals (CIs). (Bland & Altman, 2000.)

In the second stage of the analysis, the education-condition interaction was addressed. Three models were fitted. In Model 1, the association between each individual health indicator and poor health was revisited separately, and sex and age were controlled for. In the second model, education was added. In the third the interaction was introduced. Results were presented and interpreted both in terms of marginal effects and in odds ratios, as in multiplicative effects. A Wald's test was conducted for all models to assess the significance of each variable.

Several post-estimation techniques were used to obtain statistics associated with the logistic regression model and to perform overall goodness-of-fit testing for each step of the analysis. Logistic regression and measuring and reporting goodness-of-fit is an ongoing debate (Allison, 2014; Archer & Lemeshow, 2006; Menard, 2009; Steyerberg, 2008). Evaluating model fit with logistic regression has no established golden standard such as R^2 with OLS regression. Several competing options are available, such as Chi-square goodness of fit tests, Hosmer-Lemeshow tests, Classification tables, Logistic regression R^2 (or pseudo-R^2s) and using an outside data set or splitting for comparisons (Hosmer Jr, Lemeshow, & Sturdivant, 2013; Menard, 2009). In addition to the trouble with substituting R^2 with a measure of choice, most goodness-of-fit tests performed by Stata are unable to account for the survey design structure. Two substitutes for R^2 were calculated: McKelvey and Zavoina's R^2 ($R^2_{M\&Z}$) and TjurR^2.

$R^2_{M\&Z}$ is an attempt to create a formula to measure model fit as the proportion of variance accounted for. It is suitable for use while using weights so it was reported here:

(1.4) $$R^2{}_{M\&Z} = \frac{Var(y^*)}{Var(y^*) + Var(e)}$$

TjurR^2 was also calculated (also known as coefficient of discrimination) as the second alternative to the tests available with Stata. Like R^2 for linear models, Tjur's suggestion has an upper bound of 1, with higher values indicating better predictive power (Tjur, 2009). These two results are simply two possible indicators of model fit and not the equivalent of R-squared Allison (see Allison, 2014) since neither $R^2_{M\&Z}$, nor TjurR^2 has been recommended to be routinely published. With that caution, both are presented here as directional alternatives.

In the final stage of the analysis, a non-linear logit decomposition was conducted. Several formulas have been suggested for nonlinear models (Fairlie 1999, Yun 2004); here the Fairlie -method and the derivative Stata *fairlie* -package were selected. The method is less generalizable than some of the other options available such as *nldecomp* (Sinning, Hahn, & Bauer, 2008) and Oaxaca with the logit command, but it suits the study question and the complex sampling design. It can be used to disentangle group differences and multiple coefficients.

The *fairlie* -program utilises iterative sampling and pairing of the observations: a sample is drawn from the two groups selected for pairwise comparisons, and the group distributions of the coefficients are matched. Decomposition is often used in the study of for instance wage gaps in minority/majority situations, where the majority is expected to more likely have a positive result in the outcome variable. In practical terms, the size of the sample drawn is determined by the size of the minority group (smaller size, lower probability of outcome) which is then matched with a sample drawn from the majority group (larger group, higher probability of outcome). Here, the group with the lesser educational attainments was selected as the reference group for each pairwise comparison. Either one of the two groups or the whole of the three groups could in fact be used as a reference group; the results would most likely vary slightly depending on this choice. The choice is more than anything a question of preference.

After the sample has been drawn and the distributions matched, the change in predicted probabilities between the groups is calculated. The program then partitions the predicted differences in the probability of poor self-rated health for the groups, producing an estimate of the health gap between the groups that is divided into differences accounted for by the different distributions of the model covariates (observed differences between groups) and difference that is unaccounted for (unobserved residual). The formula is built on the Blinder-Oaxaca regression decomposition and can be written as (Fairlie, 2005):

$$\bar{Y}^L - \bar{Y}^H = \left[\sum_{i=1}^{N^L} \frac{F(X_i^L \hat{\beta}^L)}{N^W} - \sum_{i=1}^{N^H} \frac{F(X_i^H \hat{\beta}^L)}{N^H} \right] + \left[\sum_{i=1}^{N^H} \frac{F(X_i^H \hat{\beta}^L)}{N^H} - \sum_{i=1}^{N^H} \frac{F(X_i^H \hat{\beta}^H)}{N^H} \right] \quad (2.1)$$

In (2.1) the *L* denotes the low probability of outcome -group or the minority (here, higher educational attainments) and *H* the high probability group or the majority (lower educational attainments). N^j is the sample size for educational group *j*. The first term in the brackets accedes to the previously mentioned portion of the health gap that can be explained with the group differences in the distribution of the health variables *X*, and the second term both the portion in the differences in the outcome *Y* related to other factors and the unobservable.

As mentioned previously, Fairlie’s method produces a detailed decomposition (Jann, 2008) that assesses the contribution of each individual coefficient to the gap. To estimate the contribution, the program sequentially replaces the covariate distribution of the selected group with that of the reference group and conducts pairwise comparisons either one coefficient or sets of coefficients at a time. This can be expressed as:

$$(2.2)\quad \frac{1}{N^H}\sum_{i=1}^{N^H} F(\hat{\alpha}^* + X_{1i}^L \hat{\beta}_1^* + X_{2i}^L \hat{\beta}_2^*) - F(\hat{\alpha}^* + X_{1i}^H \hat{\beta}_1^* + X_{2i}^L \hat{\beta}_2^*).$$

Fairlie (2005) suggests that since the obtained estimates depend on the order in which each individual variable is entered, randomly drawn replicate samples combined with randomizing the variable order are advisable to average the results across all possible combinations. In the decomposition reported here 1000 random subsamples was used to calculate the means. This way correlations across the characteristics are not lost and path-dependency is addressed. As with the choice in reference group, the choice in the number or drawn samples has no established standard. Fairlie himself made cautious suggestions, which have been commented on in Statistical papers that go beyond the scope of this Thesis.

7. Results

7.1 Descriptive analysis

The majority of the 5586 subjects were female (table 1). The respondent mean age was 51.2 ranging from 30 to 93 years. The largest educational group was basic education. Reporting poor or very poor health was somewhat rare.

As shown in Table 1, since hypertension was identified using current medical recommendations, nearly half of the participants received the diagnosis. Physical activity levels falling below recommendations, high cholesterol levels and musculoskeletal disorders were also frequent, while a somewhat small number of respondents received a positive diagnosis for cardiovascular diseases, mental health disorders or diabetes.

Table 1. Sample Characteristics, Health 2000. Unweighted.

		Count	%
Poor health		526	9.4
Sex	Male	2548	45.6
	Female	3083	54.4
Age	30 - 44	1880	33.7
	45 - 64	2584	46.3
	65 - 100	1122	20.9
Education	Basic	1480	36.5
	Upper secondary	1347	33.2
	Higher	1228	30.3
Health behaviour	Daily smoker	1226	22.0
	Low physical activity	2124	38.0
	Low vegetable consumption	1053	18.9
Clinical	Obese	1233	22.1
	High cholesterol	1445	25.9
	Hypertension	2584	46.3
	Diabetes	391	7.0
	Musculoskeletal disorders	2153	38.5
	Cardiovascular diseases	609	10.9
Functional	Subjective troubles	602	10.8
	Low ability test score	892	16.0
	Clinician: limitations	1207	21.6
Symptoms	High somatisation	549	9.8
Mental health	Diagnosed disorder	696	12.5
Total		5586	100.00

Table 2 presents the distribution of health ratings by the background variables. Approximately 91 % of the respondents rated their health as very good, good, or average and only a little over 9 % as poor or very poor. A larger percentage of men than of women estimated their health as poor or very poor, although the mean of self-rated health was lower for women.

Table 2. Prevalence of the explanatory variables by self-rated health, unweighted.

		Self-rated health: Very good, good, average n	%	Poor, very poor n	%	Total n	Total %
Total		5060	90,6	526	9.4	5586	100
Sex	Male	2295	90.1	253	9.9	2548	100
	Female	2765	91.0	273	9.0	3083	100
Age	30-44	1828	97.2	52	2.8	1880	100
	45-64	2322	89.9	262	10.1	2584	100
	64 - 100	910	81.1	212	18.9	1122	100
Education	Basic	1693	83.8	212	18.9	2021	100
	Upper secondary	1735	93.1	129	6.9	1864	100
	Higher education	1632	95.9	69	4.1	1701	100
Health behaviour	Daily smoker	1084	88.4	142	11.6	1226	100
	Does not smoke daily	3976	91.2	384	8.8	4360	100
	Low physical activity	1882	88.6	242	11.4	2124	100
	Activity average or above	3178	91.8	284	8.2	3462	100
	Low vegetable consumption	907	86.1	146	13.9	1053	100
	Recommended consumption	4153	91.6	380	8,4	4533	100
Clinical	Obese	1050	85.2	183	14.8	1233	100
	Not obese	4010	92.1	343	7.9	4353	100
	High cholesterol	1283	88.8	162	11.2	1445	100
	Cholesterol average or below	3777	91.2	364	8.8	4141	100
	Hypertensions	2244	86.8	340	13.2	2584	100
	No hypertension	2816	93.8	186	6.2	3002	100
	Diabetes	301	77.0	90	23.0	391	100
	No diabetes	4759	91.6	436	8.4	5195	100
	Musculosceletal disorders	1769	82.2	384	17.8	2153	100
	No musculosceletal disorders	3291	95.9	142	4.1	3433	100
	Cardiovascuolar disorders	430	70.6	179	29.4	609	100
	No cardiovascuolar disorders	4630	93.0	347	7.0	4977	100
Functional	Subjective troubles	389	64.6	213	35.4	602	100
	No troubles	4671	93.7	313	6.3	4984	100
	Low ability test scores	727	81.5	165	18.5	892	100
	Average or high test scores	4333	92.3	361	7.7	4694	100
	Clinician: limitations	853	70.7	354	29.3	1207	100
	Clinician: no limitations	4207	96.1	172	3.9	4379	100
Somatic	High somatisation	290	52.8	259	47.2	549	100
	Average or low somatisation	4770	94.7	267	5.3	5037	100
Mental health	Diagnosed disorder	562	80.8	134	19.3	696	100
	No diagnosed disorder	4498	92.0	391	8.0	4890	100

Below-average health ratings also increased with age. While here the distribution is only presented for three separate age groups, for the persons over the age of 80 self-rated health is skewed towards “Good” and “Quite good” as is typical in large-scale survey studies. A notable percentage of those with a comprehensive level education (16 %) rated their health as poor, while the most educated seldom did so. Poor health ratings also seemed more common among those respondents with a medical condition such as a musculoskeletal disorder, cardiovascular disease, or diabetes. Mental health disorders, functional abilities – both self-assessed performance in conducting daily tasks and as abilities assessed by tests or a clinician – and somatic complaints also seem associated with self-rated health. However, Table 2 only presents basic descriptive statistics and should not be used as the base for drawing conclusions.

All the health indicators were correlated with poor health, and almost all with at least one of the other indicators. For instance, subjective functional abilities were positively correlated with clinically determined musculoskeletal disorders and BMI for instance negatively with daily smoking. All the statistically significant correlations were of low degree (<=+ .29, in most cases <.10), except for clinician’s estimation of functional ability which was moderately correlated with cardiovascular diseases, musculoskeletal disorders, and subjective evaluations of functional troubles. For a detailed table of correlations, please see Appendix 4.

Table 3 presents the distribution of each of the health indicators by education, adjusted for age and sex. Those with higher education less often rated their health as poor than those with a basic education. As expected in the light of previous studies, the prevalence of most of the potentially harmful health behaviours, medical conditions, indicators of functional health and biomarkers related to poor health is higher among those with the lowest educational attainments, while the most educated tend to fare the best by most of the measures. Daily smoking, partaking in very little physical activity and exercise, fresh vegetable consumption, obesity and all the functional measures suggest that the least educated may differ from those with a higher education. In some instances, like daily smoking, upper secondary education may differ from both the less and the more educated groups, while for instance with exercise it lines up with comprehensive education. The only health measurement by which none of the groups differed was mental health: when age and sex were adjusted for, being diagnosed with a disorder could possibly be equally likely in all educational groups.

Table 3. Age- and sex adjusted prevalence (%) of health indicators and conditions by educational level.

		Education								
		Basic			Upper secondary			Higher		
		%	95 % CI		%	95 % CI		%	95 % CI	
Poor self-rated health		11.0	9.5	12.5	6.9	5.7	8.1	4.7	3.6	5.7
Health behaviours	Daily smoking	29.3	27.1	31.6	20.9	19.0	22.8	11.7	10.1	13.3
	Physical inactivity	39.9	37.5	42.2	39.1	36.9	41.4	34.1	31.8	36.4
	Low vegetable consumption	26.7	24.5	28.8	17.7	16.0	19.5	9.8	8.3	11.2
Clinical	Obesity	26.9	24.5	28.8	21.1	19.2	23.0	16.5	14.6	18.3
	High cholesterol	28.3	26.1	30.5	26.2	24.2	28.3	20.5	18.4	22.5
	Hypertension	51.0	48.4	53.5	46.5	43.9	49.0	39.8	37.2	42.5
	Diabetes	6.3	5.2	7.4	5.4	4.4	6.5	4.2	3.2	5.2
	Musculoskeletal disorders	41.9	39.5	44.3	39.2	36.8	41.5	30.9	28.5	33.3
	Cardiovascular disease	5.3	4.4	6.3	4.3	3.4	5.2	3.2	2.4	4.0
Functional	Subjective troubles	13.9	12.2	15.6	8.4	7.1	9.7	6.1	4.9	7.2
	Low ability test score	19.2	17.3	21.1	14.9	13.3	16.6	13.4	11.8	15.0
	Clinician: limitations	22.0	20.0	24.1	18.3	16.4	20.2	10.9	9.3	12.5
Symptoms	High somatisation	12.6	10.9	14.3	7.9	6.6	9.1	3.9	2.9	4.9
Mental health	Diagnosed disorder	12.2	10.6	13.8	11.9	10.4	13.4	12.1	10.5	13.8

7.2 Logistic regression analysis

All the physical condition -variables (model 1, table 4) except for cholesterol levels exceeding recommendations and hypertension were first found to be associated with poor self-rated health when age and sex were controlled for. The associations were strongest for musculoskeletal disorders and cardiovascular diseases, all three of the measures of functional health, mental disorders, and somatic complaints. For example, the odds of poor self-rated health increased approximately 3,6 -fold for those respondents who scored low in ability tests against those who did not.

Table 4. Logistic regression, the association between the selected health indicators and poor self-rated health (in OR).

Health dimension		Model 1.			Model 2.			Model 3.		
		OR	95 % CL		OR	95 % CL		OR	95 % CL	
Health behaviour	Daily smoker	2.15**	1.73	2.70	1.63**	1.25	2.14	1.58**	1.21	2.08
	Physically inactive	1.53**	1.26	1.84	1.21	0.95	1.52	1.20	0.95	1.52
	Low vegetable consumption	1.63**	1.31	2.03	1.07	0.82	1.40	1.04	0.79	1.36
Clinical health	Obesity	1.76**	1.44	2.15	1.03	0.79	1.35	1.01	0.78	1.32
	High cholesterol	1.04	0.85	1.29	1.04	0.81	1.34	1.03	0.81	1.33
	Hypertensions	1.23	0.99	1.53	1.00	0.77	1.29	0.99	0.76	1.27
	Diabetes	1.92**	1.46	2.53	1.24	0.87	1.76	1.23	0.87	1.75
	Musculoskeletal disorder	3.77**	3.05	4.68	1.72**	1.32	2.26	1.71**	1.31	2.24
	Cardiovascular disease	2.80**	2.19	3.57	1.78**	1.32	2.40	1.75**	9	2.37
Functional health	Subjective troubles	7.59**	6.10	9.45	2.52**	1.89	3.35	2.50**	1.88	3.32
	Low ability test score	3.64**	2.94	4.52	1.70**	1.29	2.24	1.69**	1.29	2.23
	Clinician: limitations	7.61**	6.09	9.51	3.01**	2.29	3.96	2.99**	2.27	3.93
Symptoms	High somatisation	12.28**	9.83	15.35	5.00**	3.85	6.50	4.87**	3.74	6.34
Mental health	Diagnosed disorder	4.05**	3.20	5.12	2.86**	2.13	3.85	2.91**	2.16	3.93
Education	Higher education							1		
	Upper secondary							1.05	0.74	1.49
	Comprehensive							1.35	0.96	1.91
$R^2_{M\&Z}/R^2_{Tjur}$					40.1/30.0			40.4/30.1		

F(9,5577)

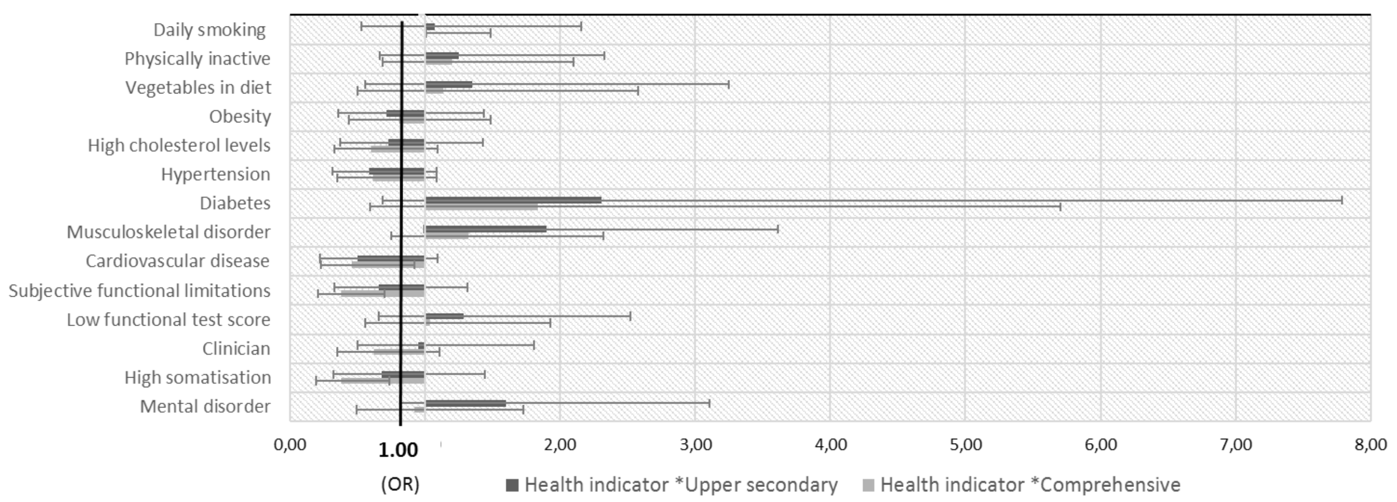
When all the health dimensions and indicators were considered simultaneously (Models 2 and 3, table 4), somatic complaints, all the functional health variables, cardiovascular disorders, diagnosed mental disorders and daily smoking retained a statistically significant association with poor SRH. This association remained when education was added (Model 3).

Even though somatic complaints indicate pain and bodily experiences that are most likely unpleasant, the functional dimension of health and some of the diagnosed diseases did not lose their association with poor self-rated health even when they were simultaneously present in models 2 and 3. Reporting subjective troubles in daily tasks and activities still more than doubled the relative risk of reporting poor health when compared in relation to baseline odds. With all the health variables present in model 3, education did not have a significant association with poor self-rated health; the association was only present when analysing single health indicators and their association with poor SRH, and excluding other health variables.

7.3 Interactions

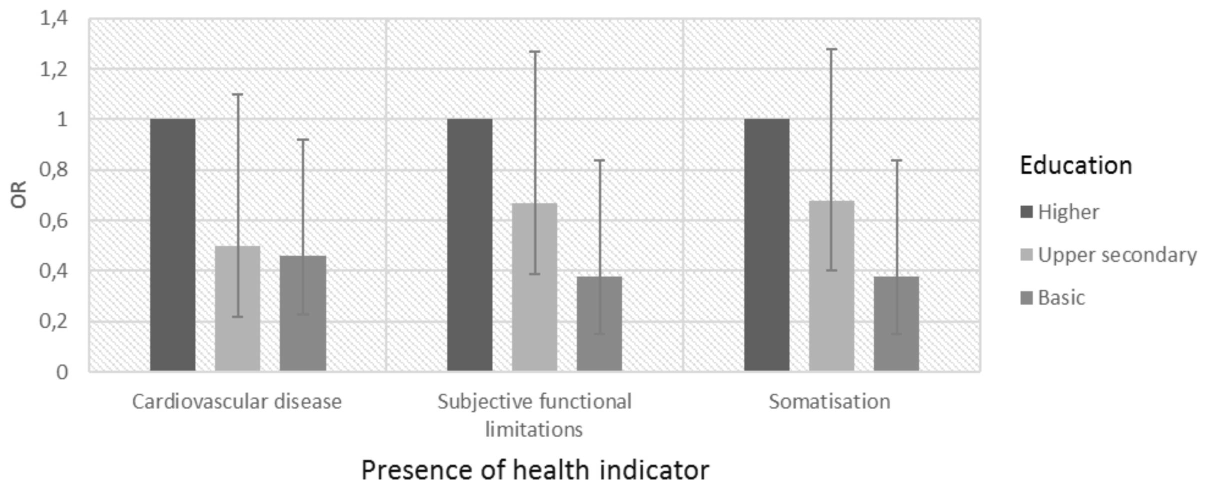
Figure 2 demonstrates the interactions by education in the association between poor self-rated health and the selected health indicators. The results are presented in odds rations with 95% confidence intervals (CI), while controlling for sex and age.

Figure 2. Health indicator * education interaction in odds ratios (with 95% CI) for poor or very poor SRH, adjusted for age and gender.



As presented in figure 2, no interactions were detected in the health behaviour indicator – self rated health -associations. An interaction was however observed in somatic symptoms, subjective functional limitations and unexpectedly, cardiovascular diseases. The direction of the effect is the same for all the previously mentioned. The health indicator was associated with poor self-rated health in all educational groups, but the odds ratios for reporting poor health in the presence of any of the conditions when compared to those without the condition were smaller for those with a basic education than for those with a higher education.

Figure 3. Health indicator * education interaction in odds ratios (with 95% CI) for poor or very poor SRH, adjusted for age and gender.



In figure 3, the above-mentioned interactions are presented in odds ratios. Odds ratios depict multiplicative effects and are as such always relative to baseline odds within their own category group. ORs therefore account for the possible differences in the likelihood reporting poor health between the categories studied. The effect of for example experiencing functional limitations in daily tasks to poor self-rated health is lesser in the case of all these three indicators for those respondents with a comprehensive level education when compared to those with higher education. The effects of, for instance, a high level of somatisation on poor health was exceedingly large, but for those with basic education it was only 0.38 of the effect size of that of the highly educated. The increase in the odds ratio that somatic complaints result in is then greater for the highly educated.

Reporting poor health was rare for those with a higher education, as was a positive outcome (diagnosis, limitation, risk factor) on almost any of the health indicators. The odds for reporting poor health for every combination of subjective functional limitations/education/somatisation with education are displayed in table 5.

Table 5. The marginal effects as the difference between the expected odds for each educational group.

	Educational level		
	Basic Margin (95% CI)	Upper secondary Margin (95% CI)	Higher Margin (95% CI)
No cardiovascular disease	0.13 (0.11-0.15)	0.06 (0.05-0.07)	0.03 (0.02-0.04)
Cardiovascular disease diagnosed	0.48 (0.37-0.58)	0.26 (0.15-0.37)	0.32 (0.14-0.50)
No subjective limitations	0.11 (0.09-0.13)	0.04 (0.03-0.05)	0.02 (0.02-0.03)
Subjective limitations	0.60 (0.47-0.74)	0.48 (0.30-0.65)	0.39 (0.22-0.56)
Average or low somatisation	0.10 (0.08-0.11)	0.04 (0.03-0.05)	0.03 (0.02-0.03)
High somatisation	0.92 (0.72-1.14)	0.74 (0.48-1.00)	0.82 (0.39-1.25)

The differences, and the differences in differences were calculated using margins: the results display how the odds in reporting change in each category of the education –variable when the health condition –variable changes from 0 to 1 (controlling for age and sex). Stata *lincom* -command was used after margins for pairwise computations. After fitting the regression model the command was used for obtaining estimates for each educational level and pairwise comparisons of the statistical significance of the group specific differences in the estimates.

The reference group baseline odds (no condition, highly educated) in the case of for instance somatic complaints were therefore .03, and so for every 100 respondents reporting average or good health with average or low somatisation and higher educational attainments, it is expected to find 3 respondent reporting poor health. For those with an upper secondary and those with basic education, the odds were 0.04 and 0.10, respectively. In other words, in the absence of the condition we expect to find a much larger proportion of those respondents reporting poor health in the basic educational group than in the higher.

Because of the educational differences at the baseline, reporting somatic complaints increased the risk of poor self-rated health 39-fold for the highly educated, 32 times over for those with secondary attainments and 35 times over for those with basic education. The corresponding volumes for subjective functional abilities were 27, 23 and 25 times over and for cardiovascular diseases 14, 12 and 11 times over, respectively.

7.4 Decomposition

Table 6 presents the detailed decomposition: not only the overall gap in reporting poor health for the educational groups, but also how much of the gap is due to differences in the covariates in the model (health indicators). As was expected, the gap in health is largest between the higher educational group and basic education, but it is also notable between secondary and basic education.

The selected health indicators and demographic variables explain virtually the whole of the difference between the groups for reporting poor health. When comparing higher to basic education, the largest portion of the gap was explained by somatization and age, since the respondent with higher education reported less pains and were younger. Other factors that contributed to the gap were clinician's estimations of subjective abilities which explained approximately 2 % of the 11 % of the gap, and subjective health limitations, cardiovascular diseases, and musculoskeletal disorders. The higher prevalence of most of the diseases and health-related problems in the group with basic education therefore, to a large extent, explains the poor SRH outcomes of the less educated. Sex, daily smoking and diagnosed mental disorders also contributed to the gap. None of the selected health indicators bridged the gap in a statistically significant level. However, comparing basic to secondary education, the sex composition and smoking habits of the groups reduced the size of the health gap between the two groups.

Table 6. Logit pairwise decomposition of educational group differences in reporting poor or very poor health.

Education	**Coefficient**
Higher education	0.042
Secondary education	0.069
Basic education	0.159

		Higher vs basic		Higher vs secondary		Basic vs secondary	
Difference			11.73		2.65		9.10
Contribution to difference		Coefficient	%	Coefficient	%	Coefficient	%
Demographic	Age	0.031**	2.63	0.021*	0.01	0,026**	1.96
	Gender	0.608**	0.34	0.157	0.06	0.393**	-0.42
Health behaviours	Daily smoking	0.424**	0.06	0.479*	0.05	0.479**	-0.24
	Low physical activity	0.152	0.00	0.276	0.03	0.173	-0.04
	Low vegetables consumption	-0.065	-0.00	0.037	0.01	0.117	0.01
Clinical	Obesity	0.053	0.00	-0.058	-0.01	0.011	0.00
	High cholesterol	-0.014	-0.01	-0.238	-0.02	-0.020	-0.00
	Blood pressure	0.081	0.12	-0.126	-0.04	-0.047	-0.06
	Diabetes	0.123	0.00	0.268	0.04	0.298	0.15
	Musculoskeletal disorders	0.437**	0.85	0.485*	0.22	0.650**	0.45
	Cardiovascular diseases	0.570**	0.95	0.993**	0.07	0.417**	0.59
Functional health	Subjective limitations	0.959**	1.35	1.020**	0.16	0.825**	0.96
	Low ability test score	0.546**	0.25	0.430	0.06	0.536**	0.11
	Clinician: limitations	1.030**	2.04	1.267**	0.44	1.078**	1.28
Symptoms	High somatisation	1.491**	2.66	1.935**	1.04	1.552**	2.64
Mental health	Diagnosed disorder	0.804**	0.02	1.459**	0.43	1.066**	0.13
Total explained		-6.490	10.72	-6.039**	2.55	-6.004	7.56

Contribution estimates are mean values of the decomposition using 1000 subsamples of basic education,
Logit estimate of decomposing the likelihood of poor self-rated health (Dichotomous variable; srh1=1, poor SRH),
∗ $p < 0.05$, ∗∗ $p < 0.01$

The gap in poor SRH between higher education and secondary education was small. Almost one third of it was explained with the differences in somatisation. Except for the results of the functional ability testing and sex, the variables that contributed to the gap corresponded those found in the previous pairwise comparisons.

8. Discussion

8.1 Comparisons with previous studies

The findings made in the first section of the analysis partially reflect the existing literature on self-rated health. A high volume of somatic complaints was found strongly associated with poor or very poor self-rated health, even when the other health dimensions were controlled for. Other significant associations were those between poor SRH and the selected measures of functional health, including tests of capability, diagnosed mental health conditions and two of the three clinically defined diseases; also, to a lesser extent, daily smoking. These results for the most part concur with previous studies. Although, some studies have also found health behaviours and several biomarkers associated with SRH, while no such association was observed here (Goldman et al., 2004; Jylhä, 2009; Sohn, 2015).

Most of the biomarkers and health behaviours showed either no association with poor SRH, or the association disappeared when the other health conditions were added to the model. The biomarkers used here have in previous studies been found associated with SRH (Dowd & Zajacova, 2010; Idler & Kasl, 1995; Idler et al., 2000). Initially, BMI indicating obesity was associated with poor SRH, but the association no longer statistically significant after the other health factors were added in Model 2.

A previous study found self-reported health behaviours associated with 3-point symmetric wording of SRH (Manderbacka et al., 1999). In the crude model stage of this study, a similar association was found for a wide range of health behaviours, but after controlling for self-reported diseases, medications, pains, and sociodemographic factors, only smoking and vegetable consumption maintained the association. In this thesis, vegetable consumption did not maintain the association, while smoking did.

Earlier findings on interaction effects have been dissonant. Some researchers have reported systematic differences in reporting styles (Down, Zajacova 2010). However, with the single exception of Onadja and colleagues (Onadja et al., 2013), all the studies that addressed other markers for health than mortality agreed, that education levels strengthened or weakened the associations, but they remained significant throughout for all the educational groups.

Similar conclusions were reached here in general, but only for three variables. Also, the variables for which an interaction was found were somewhat surprising. As stated in the previous chapter, interaction effects were found for subjective experience of functional limitations in everyday activities, high number of somatic complaints and diagnosed cardiovascular disease. The three health indicators all represented different health dimensions: those of bodily symptoms, functional health, and clinical health. Delpierre, (2012) based on the American NHANES -data and the asymmetric SAH wording, also found the highly educated more affected by subjective functional troubles. Meanwhile, the rest of the observations do not repeat the conclusions of previous studies. Education did not systematically affect the association between the health indicators and SRH for health behaviours as discovered by Sohn (2015), or health biomarkers as Dowd and Zajacova (2007) suspected. Also, Sohn (2015) earlier hypothesised, that when a diagnosis from a medical professional is combined with some suitable symptoms of an illness, this combination should have an association with SRH that is not moderated by education. The interaction discovered for CVD challenges the notion.

Onadja and colleagues (Onadja et al., 2013) discovered that the association between their measure of functional ability and self-rated health was fully explained away in the highest educational group by self-reported illnesses and diseases. The items did not measure ability in everyday tasks as was done in the Health 2000 -study, but more general abilities such as vision, worded as follow: “Do you have difficulty seeing, even if wearing glasses?”.

Not all the clinical diseases were associated with poor health. For example, diabetes was not. The Health 2000 data does not separate between different types of diabetes. To explain this, a case in point is the discovery related to cardiovascular and ischemic heart diseases. Non-ischemic cardiovascular diseases and their processes are often manifested as various symptoms. A study found a steadily decreasing trajectory of SRH prior to death from non-ischemic diseases. Ischemic cardiovascular diseases on the other hand are linked to clustering of such biomarkers as high blood pressure and sugar levels, obesity, and atherosclerosis prior to death from said infliction, while the death itself is sudden. Mirroring this, the trajectory of SRH prior to the latter cause of death is a steady period of sub optimal health instead of the slow decline. From this and other similar findings it follows that perhaps even a medical condition is only included in SRH if it negatively affects the everyday life of the person. This could partially explain the findings concerning diabetes. If the condition becomes a part of the normal, everyday life of the patient and causes little physical pain, it may not become included in the assessment. (Stenholm et al., 2016.)

8.2 Education as a confounder

An interaction effect was found for three of the health indicators. Some researchers have found a more systematic effect, which they have attempted to explain. Differences in health literacy and health monitoring habits are among the most common explanations. For example, Sohn (2015) hypothesises that educated individuals may have a higher level of health literacy and may benefit from more regular self-monitoring and health check-ups.

Health knowledge and awareness is a particularly popular explanation for the educational differences in the association between health behaviours and SRH. Sohn's additional study questions concerned the amount of knowledge the educational groups had of their health indicators. In working-age US population, higher education increased the likelihood that the individual had had a blood pressure and cholesterol levels check at some point during the 12 months preceding the study. Even when the existence of diabetes was controlled for, a higher educational level was still associated with more accurate blood-pressure level self-assessments. Furthermore, those with higher education were also more likely to correctly report themselves as clinically overweight when their BMI exceeded or equalled 25. (Sohn 2015.)

Delpierre (2009) assumed that the awareness of the consequences of health risks may increase with education. Consequently, more educated individuals may be more aware of health risk behaviours and include, beforehand, the expected future ramifications of their health behaviours in their current health self-ratings. In an earlier study, Goldman and companions (2004) hypothesised along similar lines on the sex difference in the SRH-biomarker –relationship. They supposed that the sex difference in knowledge of health risks and the perception of the personal likelihood of a negative health outcome would explain the differing health assessment profiles.

Disparities in access to health care and health knowledge may play a significant part in the modifying effect of education on SRH (Goldman et al., 2004; Sohn, 2015). With higher education come both a better understanding of one's physical state (more likely to have health examined), and the significance of these biological indicators of health (more likely to recognise risk factors). Those with higher educational attainments may be both more knowledgeable of different measurable indicators of their health status and more likely to internalize these measures (Sohn 2015).

Firstly, in the case of this study, self-rated health data was gathered before the medical examination, and thus it is possible that not all respondents were aware of having some of the conditions and diseases. For instance, it is quite plausible that some of them might have been unaware of being hypertensive, obese, or perhaps even diabetic. Some of the conditions are less likely to go unnoticed. For example, an interaction effect was found for CVDs. But since here cardiovascular diseases consisted of heart infarct, heart failure and stroke, it is less plausible that the participants were unaware of them, than in the case of for instance hypertension.

Secondly, no interaction effect was found for the health behaviours. Accordingly, the study does not support the earlier findings that the educational groups, for one reason or the other, differ in their inclusion of health behaviours in their health self-assessments. Following Manderbacka and Martikainen (1999), this would seem sensible: the association between some of the health behaviours (namely smoking and vegetable consumption) and SRH in fact needs to be mediated through specific health problems and symptoms. For instance, a regular smoker may not have been diagnosed with chronic bronchitis, but still experience respiratory or other difficulties that are then reflected on the SRH even though the condition is still pre-clinical.

The causality in the health-SRH association should therefore be bidirectional. Poor health and bad health-related lifestyle choices may lead to poor health expectations, and the poor expectation lessen the motivation for change which may in time negatively impact the health state. (Karisto, 1984, p. 27.) Manderbacka and Martikainen (1999) add a time dimensions. Per their study, it is still unclear whether factors such as health behaviours affect SRH independently, or whether unhealthy behaviours lead first to bad health which then reflects in SRH.

Yet, is must be noted that material inequalities may also function as determinants of health behaviours. Low educational status is also associated with lower income, which may in turn limit the behavioural choices available to the individual, such as daily smoking and alcohol consumption (Martikainen et al., 2004).

Bago D'uva and colleagues (2007) note that differences in health literacy and conceptions of illness are probably less varied in developed countries. Social inequalities and health knowledge may together affect the interaction. Countries, such as the Nordics, with a low level of inequalities and universal health care and education, only the very extremes of the educational distribution would exhibit differences in areas such as health literacy. In Finland, the welfare state would function to level the differences in health knowledge and expectations. Accordingly, the strongest interaction effects have been discovered in the United States (Dowd & Zajacova, 2010; Sohn, 2015).

Following this, differences in knowledge of health status or health risks might not at least fully explain the observed interaction effects. Another popular explanation for reporting heterogeneity is founded in the respondents' views and experiences of what is normal health. Per the social comparison theory, individuals make comparisons in relation to their experiences of their peers (Schnittker, 2005). For instance, highly educated and wealthy respondents would therefore not compare themselves to the population in general, but their healthy peers instead. Viewing great health as normal would result in a lower threshold for poor health, and a more restrictive threshold for great health. Groups with lower average health would accordingly be prone to rating their health up, since the cut-off point for great health would be lowered by their experience of the average health status of their peers. Both Delpierre (2009) and Sohn (2015) agree that in any case, the existence of health problems is more damaging to the experience of self-rated health for those with a higher level of education. This was also found to be the case for some of the indicators utilised in this study.

Some theorists have challenged the implications of the association between high education or socioeconomic status and SRH. Per the wishful thinking –theory, a higher socioeconomic status or education may lead to more positive health self-assessments simply because the well-off respondents believe they should be healthy (Iburg, Salomon, Tandon, & Murray, 2001). Therefore, education predicts high SRH, but not necessarily good health. This would lead to overestimating existing socio-economic health differences.

In the light of this data, the wishful thinking –explanation, however, does not appear to get support. The highly educated have both better health and higher self-rated health. Yet, it is possible that the violated health expectations explain why health problems are more damaging to the self-rated health of the highly educated. The most plausible explanation appears to be that of the differences in peer health experiences, which then leads to differing ideas of normal and expectable health between the groups.

8.3 Strengths of the study

The population-based study design, large sample sizes and high participation rates are all important strengths of this study. Also, there is a wide variety of different types of indicators for health which makes it possible to compare SRH to a wide definition of health. Objective measures of functional capacity and health were included.

A strength of this study is also the rich data. Gaining a “perfect” understanding of the health status of an individual is difficult, or according to some even impossible. An examination by a trained physician such as conducted in Health 2000 provides a good overall picture. The provides many rare opportunities: SRH is compared to a wide selection of health indicators derived from clinical evaluations, functional performance, and laboratory tests, as well as self-reported items, as suggested in previous studies.

Here, the objective health data were collected by medical professionals from a random sample data. Some of the previous studies compared SRH to lists of self-reported illnesses instead. Per Goldman (2004), and Bjorner and colleagues (1996, p. 39) comparing self-reported illnesses to self-reported health is problematic because of issues related to response style, poor recall, understanding of the medical condition, or the reluctance to confide in the interviewer. Preferably, conditions reported by a trained medical professional should be used instead (Bjorner et.al. 1996, p. 39; Thomas and

Frankenberg 2002; Dowd and Zajacova 2010). Since the likelihood of visiting a skilled medical professional is socially patterned, a medical examination conducted for a random sample is the recommendable option.

The medical and health sciences have traditionally given great value to the objective and the observable, such as the above-mentioned indicators of disease. Purely objective health would also, however, be a poor single measuring tool against which to test SRH, since not every dimension like somatisation can be measured in a way that could be called objective. The Health 2000 data also included a section that was used to represent the functional health -dimension of health combining both objective and subjective measures and an interview on somatic complaints and unspecific symptoms. (Karisto, 1984, p. 24-25.)

In this study, many of the health indicators correspond to the negative definition of health. In its defense, this approach is both practical, grounded in real troubles of the material every-day life and is often used and considered feasible in policy making (Karisto, 1984, p. 8-9). Poor health is not just something possibly correlated with illness, it is also an indicator of negative emotional states such as stress and helplessness, as Karisto (1984, p. 68-69) reminds us.

8.3 Limitations

Firstly, dichotomisation, of the outcome variable and the health indicators, can be considered a limitation. While dichotomising categorical or continuous variables has been noted to inevitably lead to some loss of information and possibly reduced efficiency (McCullagh, 1980), dichotomisation was still used here. One of the reasons for this was simply the widespread use of SRH as a binary variable. Also, had the variable not been dichotomised, combining at least the poor and very poor –categories would have been inevitable to guarantee a sufficient minimum amount of responses in each health category (Garson, 2012).

Not all studies agree on the outcome variable cut-off point utilised here. And yet, only a few studies have addressed the effects of SRH cut-off point for the study results. Here, the analyses were also conducted using the other widely used cut-off point for

SRH as an outcome; average, poor, and very poor (less-than-good SRH) health. Changing the cut-off point did not radically alter the results, though an additional interaction effect was found also for BMI. Since many studies, including this one, dichotomise self-rated health, testing the cut-off point is a topic for future studies.

A study (Manor, Matthews, & Power, 2000) also tested the effect of the dichotomisation of 4-point Likert scale self-rated health on a question on the effects of SES. Logistic regression did produce similar results as others methods not based on binary response variables. The differences in power and efficiency were small.

Having to make choices for category cutting-points for each health indicator was also unavoidable. The choices in this study were made in accordance with previous research and following current, accepted medical guidelines and recommendations. Several different cut-points are, however, also used for several of the health indicators used here. BMI provides a good example. Categorising the measure into dummy groups hid some of the original continuum inherent to the measure. If “normal weight” and “slightly overweight” are piled together, some information may be lost. Categories draw artificial lines, for example when separating someone with a BMI of 24 kg/m^2 from someone with that of 25 kg/m^2. Such choices make the cut points manipulative. Still, since a BMI slightly above the recommended normal upper limit may not pose a threat to health (for example, athletes often exceed this limit) and a BMI exceeding 30 is already a notable health risk, the choice was justifiable. For the purposes of this study it is perhaps enough to keep these limitations in mind. (McDowell 2006, p. 32.)

In addition to the cut-off points, the results of this type of study and the heterogeneity patterns discovered are always sensitive to the health indicators and measures that were selected. The indicators used in this study have however been used and studied extensively (Shmueli, 2003).

The survey health indicators themselves have their limitations. For instance, Bjorner (1996, p. 55) reminds us that questions concerning somatic symptoms such as those collected with the SLC-90 often overlap with general health. Further examples are the survey items that measure physical activity. The questions instructed the respondent to evaluate and include “exercise” that would at least leave him/her slightly short of breath, and his/her method of commuting to work. Household chores and physical labour were not included in the instructions. It is possible that measuring activity in this manner favours the more educated, who are more likely to participate in this type of activity. The measure is blind to for instance demanding physical labour or chores.

A study found self-rated health associated with physical activity related energy expenditures in almost all European countries, though Finland was found one of the four exceptions (Abu-Omar, Rütten, & Robine, 2004). Including items on other types of physical activities than exercise might have produced a more inclusive indicator.

One of the strengths of this study is the inclusion of both objective and subjective health indicators. Especially the use of clinical health evaluations instead of self-reported disease has been recommended in several previous studies. Dowd & Zajacova (2010) note that biomarker data collected from a random sample is less prone to systematic reporting heterogeneity than, for instance, many of the subjective indicators of health. Thomas and Frankenberg (2002b, p. 397) however add that this may not be the case for some of the assessments performed by medical professionals. If the assessment requires some form of participation from the study subject, as for instance the standard puff-test, the subject’s individual characteristics may influence the participation. For instance, subjective performance tests such as were conducted here require the subjects’ effort and interaction with the medical personnel and are thus dependent on willingness to perform.

The SRH data was acquired at the very beginning of the home interview and prior to the medical exam section of the study. The medical examination was conducted quickly following the home interview (usually within a few weeks) and therefore substantial changes in the health status of the respondents are not highly likely. However, changes in the conditions measured in the latter section of the study are still possible.

Gaining a perfect understanding of the health status of an individual is difficult at best. An examination by a trained physician such as conducted in Health 2000 provides a conclusive overall picture for quantitative study purposes. The Health 2000 data offers a wide definition of health and provides many rare opportunities, but still it is impossible to cover health as a whole. What about variables such as the family medical history (Thomas & Frankenberg, 2002b)? Several interesting factors and determinants were excluded from this study.

Personal health believes and attitudes were not addressed in this study. Several researchers have suggested including so-called anchoring vignettes to survey studies on self-rated health. Here, anchoring vignettes refer to separate survey study components, that typically comprise of descriptions of health-related situations that measure the respondent's health expectations and beliefs. The vignettes are then used to address the assumed incompatibilities between either individuals, or population sub groups. Vignettes however are not typically added to general health surveys so they would require new data collection specifically for that purpose (Lindeboom & Van Doorslaer, 2004; Salomon et al., 2004). Attempting to control for health believes using this data goes beyond the scope of this Thesis.

The SRH is a subjective measure, and as such possibly affected by the personality of the respondent. As with health beliefs and attitudes, controlling for personality factors was not feasible. Furthermore, if the affecting personality factors are not typical for a single educational group this should not be considered too problematic.

To what degree the individuals concurred with the diagnoses they received in the clinical estimation was also deemed beyond the scope of this Thesis. The Health 2000 -study does in fact include the typical binary survey item: "Do you have some long-standing illness or disability", which was controlled for in the early stages of exploratory model building. However, since it is a binary variable, it is unclear to which disease or conditions the respondents were referring to, and whether it was included in this Thesis. In the data, there were several participants who experienced themselves as having some illness or disability, but were declared healthy in the medical examination. Vice versa, there were some who felt healthy, but had a diagnosed medical condition. The confounder was ultimately omitted.

In the exploratory stages of building this study, geographical location and the urban versus rural –divide were accounted for using the standard statistical grouping by Statistics Finland (here based on the 2000 Population Census: urban, densely populated, rural) and the University hospital districts. Based on previous studies (Bago d'Uva et al., 2008) geographic location was expected to be associated with self-rated health due to health expectations and the accessibility and quality of health care services. Since the association between SRH and geographical location disappeared as soon as any one of the confounding health indicators, age or education was added to the regression model, location was omitted. No interaction effect was observed.

Knowledge of family health history was not included in this study; along with other factors such as unhealthy lifestyles and subclinical conditions it has been suggested to affect health self-ratings (Idler & Kasl, 1991). Genetic and environmental factors were also deemed outside the scope of this Thesis, even though admittedly both are likely to be relevant to both health and health self-ratings. Also, disease severity is one possible confounder not addressed here. The modelling strategy is able to some extent account for the overlapping of disease, but the accumulation of health problems and its relation to SRH deserve more attention.

Heavy drinking was left outside this study because it showed no association with poor health. The results mirror previous studies (Manderbacka, Lahelma and Martikainen 1998; Lahelma et.al. 2010). However, because alcohol consumption was measured using self-reported amount of consumption turned into grams of alcohol consumed, a person who consumed a small glass of red wine at dinner every night appeared in the data the same as the person who consumes two bottles on one sitting.

Also noteworthy are Quesnel–Vallée's (2007) previous conclusions on the missing association between heavy drinking and SRH. Since binge drinking raises risks to health from external causes (such as accidents) that are by nature unpredictable and not necessarily preceded by poor health, this type of drinking habits might not be reflected in SRH at all. A rivalling explanation is selection: those in very poor health might not be able to consume alcohol regularly. Targeting instead those respondents that consumed very little alcohol was also questionable. Some of the current non-users might have had a history of heavy drinking or had quit consuming alcohol due to adverse health outcomes. The only group that could have been separated from this data was those respondents who had never been regular consumers at all.

Even with the sizable data, the combination of poor self-rated health with on any of the indicators of health and higher education was rare. For an example, diabetes was quite rare among the highly educated and out of those few respondents who did have it, only 6 % rated their health as poor. Thus, the final cell sizes were often small. The equivalent ratio of poor SRH for those with comprehensive education was 30 %.

Socially patterned study participation rates are a persistent problem with survey studies (Koskinen et al. 2008, p. 132-143). To make the participants data fully comparable, cases with missing data were excluded. A missing data analysis was conducted on the remaining data, but the fact remains some losses were inevitable. At least the proportion of the very old was possibly reduced because of the cut as was the proportional size of the "very poor" self-rated health group, and the proportion of those with a cardiovascular disease. In the case of most of the health variables, the variation in group proportional sizes either remained the same or only changed by a few percentage points, and the losses to data quality were considered acceptable.

8.4 Conclusions and implications for future studies

This study seems to support the notion that the Finnish population has a shared understanding of poor health. Poor SRH seems a good overall measure of poor health, although the study found some indication of the less educated being more prone to ignore pains, functional limitations, and cardiovascular diseases in their health self-assessments, when compared to the most educated group. Also, the decomposition implies that the differences in SRH mostly consist of differences in the presence of the selected health indicators. These negative conditions were also more prevalent among the basic education group.

Yet, average, or good health may be harder to agree upon. Out of curiosity, I also ran the *fairlie* decomposition with the outcome of reporting *very good health* (running it for the outcome of not reporting poor health would have included approximately 90

% of the respondents). The results and their meaning go beyond the scope of this study, but for instance the gap between the highest and the lowest educational group in very good health was almost three times as large as it was in poor health. Additionally, only one third of the gap was explained by the coefficients chosen for this study, compared to them almost explaining the whole gap with poor health as the outcome.

Further studies are needed to both target the full of range of SRH, as well as excellent self-rated health to study whether the criteria for very good health is different for the three educational groups. Good health is not usually addressed as an urgent societal issue in the same way that lack thereof is, for obvious reasons. Still, if acceptable health is different for those who are well-off, those up the socio-economic ladder should likely both ask and receive more than the deprived possibly contributing to the health gap (Karisto, 1984, p. 24-25).

In addition to extending the study question to include excellent health, the changes and differences in educational group health beliefs and expectancies have, to my knowledge, not been studied in Finland. A life that revolves around health can be equally anxiety-inducing as habitually engaging in poor health behaviours. Constantly tightening the criteria for acceptable health may be counterproductive for all the educational groups: it may essentially favour the most educated, but also cause them stress and worry. (Karisto 1984, p. 77-88.)

Awareness of disease is one possible topic for future studies. A Norwegian study (Jorgensen et al., 2015) tested the associations of laboratory results indicating diabetes mellitus, hypertension or hypothyroidism and self-reports of these diseases in a survey to 4-point SRH. In the cases when clinical tests indicated one of the diseases but the study participant did not report accordingly, he or she was classified as unknowingly having the disease. Known hypothyroidism (women only), diabetes and hypertensions were found to increase the odds ratio for reporting poor self-rated health at the 11-year follow-up. SRH was measured using a dichotomised 4-point scale. No such association was found for unknown disease. The study question concerns the negative health-reporting effects of disease labelling, but the results may also indicate that some health states with mild or no symptoms only become reflected in SRH after being diagnosed by a clinician. Attempting to control for the effects of disease awareness with the Health 2000 data could be an interesting question.

The Health 2000 data includes a question on whether the respondent considered that he/she had some chronic or long-term disability or illness that lessened his/her ability to work or function. When compared to the clinical assessments, quite a few of the respondents had diseases that they at least did not report under the question. Vice versa, hundreds of respondents reported having illnesses even though judging by the indicators included in this study none could be found. The internalization of health information is likely to affect how the individual sees his or her health.

Narrowing off inequalities in health has been on the target list of Finnish health politics since 1986. Several high-quality studies have confirmed the existence of a persistent socioeconomic ingredient in specific-cause mortality (see Jylhä 2009; Tarkiainen et al 2012) and self-rated health (Heistaro et al., 1996; Helakorpi et al., 1999; Rahkonen et al., 2004) in Finland. Reducing health inequalities is one of the five cornerstones of the Finnish Ministry of Social Affairs and Health's program for the reign of the current Government, appointed in 2015. Self-rated health has been lifted on a pedestal in the Health 2015 public health program as an indicator for population health, and as a target in itself. Objective seven of the program states that the self-rated health of the Finnish population must either maintain its current level or improve (Ministry of Social Affairs and Health, 2001).

Poor health is often not just addressed as a private problem, but as a collective loss of a functioning, working member of the society. Health is personal, but it mixes with the good of the society and is followed by responsibilities laid on the individual to stay healthy. Already in 1980's, Karisto (1984, p. 73-76) saw the weight of the common good falling on the individual and personal health tied with the wealth and competitive capabilities of the nation.

The implications of including self-rated health in the objectives of Health 2015 can be viewed as respectful to the health beliefs and experiences of the individual. SRH has been suggested to include an element of health-related quality of life and satisfaction with health status. The objective however targets maintaining or improving SRH measures, instead of reducing disparities in self-rated health. This may be an opportunity lost for advancing equality.

References

Abu-Omar, K., Rütten, A., & Robine, J. (2004). Self-rated health and physical activity in the european union. *Sozial-Und Präventivmedizin, 49*(4), 235-242.

Adams, J., & White, M. (2006). Is the disease risk associated with good self-reported health constant across the socio-economic spectrum? *Public Health, 120*(1), 70-75.

Alberti, K. G. M. M., & Zimmet, P. F. (1998). Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: Diagnosis and classification of diabetes mellitus. provisional report of a WHO consultation. *Diabetic Medicine, 15*(7), 539-553.

Allison, P. (2014). Measures of fit for logistic regression. *SAS Support,* (Paper 1485-2014), 25.07.2016.

Andersson, S., Ekman, I., Friberg, F., Daka, B., Lindblad, U., & Larsson, C. A. (2013). The association between self-rated health and impaired glucose tolerance in swedish adults: A cross-sectional study. *Scandinavian Journal of Primary Health Care, 31*(2), 111-118.

Anson, O., Paran, E., Neumann, L., & Chernichovsky, D. (1993). Gender differences in health perceptions and their predictors. *Social Science & Medicine, 36*(4), 419-427.

Antonovsky, A. (1987). *Unraveling the mystery of health: How people manage stress and stay well.* Hoboken, NJ: Jossey-Bass.

Archer, K. J., & Lemeshow, S. (2006). Goodness-of-fit test for a logistic regression model fitted using survey sample data. *Stata Journal, 6*(1), 97-105.

Aromaa, A., Heliövaara, M., Impivaara, O., Knekt, P., Maatela, J., Joukamaa, M., . . . Mälkiä, E. (1989). *Terveys, toimintakyky ja hoidontarve suomessa. mini-suomi-*

terveystutkimuksen perustulokset. Helsinki: Kansaneläkelaitoksen sosiaaliturvan tutkimuslaitos, Kansaneläkelaitoksen kuntoutustutkimuskeskus.

Aromaa, A., & Koskinen, S. (2004). *Health and functional capacity in finland : Baseline results of the health 2000 health examination survey.* Helsinki: National Public Health Institute.

Bago d'Uva, T., Van Doorslaer, E., Lindeboom, M., & O'Donnell, O. (2008). Does reporting heterogeneity bias the measurement of health disparities? *Health Economics, 17*(3), 351-375.

Bailis, D. S., Segall, A., & Chipperfield, J. G. (2003). Two views of self-rated general health status. *Social Science & Medicine (1982), 56*(2), 203-217.

Balog, J. E. (2005). The meaning of health. *Journal of Health Education, 36*(5), 266-273.

Baron-Epel, O., & Kaplan, G. (2001). General subjective health status or age-related subjective health status: Does it make a difference? *Social Science & Medicine (1982), 53*(10), 1373-1381.

Baron-Epel, O., Kaplan, G., Haviv-Messika, A., Tarabeia, J., Green, M. S., & Kaluski, D. N. (2005). Self-reported health as a cultural health determinant in arab and jewish israelis MABAT--national health and nutrition survey 1999-2001. *Social Science & Medicine (1982), 61*(6), 1256-1266.

Barsky, A. J. (1988). The paradox of health. *New England Journal of Medicine, 318*(7), 414-418.

Barsky, A. J., Cleary, P. D., & Klerman, G. L. (1992). Determinants of perceived health status of medical outpatients. *Social Science & Medicine, 34*(10), 1147-1154.

Bath, P. A. (2003). Differences between older men and women in the self-rated health–mortality relationship. *The Gerontologist, 43*(3), 387-395.

Benjamins, M. R., Hummer, R. A., Eberstein, I. W., & Nam, C. B. (2004). Self-reported health and adult mortality risk: An analysis of cause-specific mortality. *Social Science & Medicine, 59*(6), 1297-1306.

Benyamini, Y., Leventhal, E. A., & Leventhal, H. (2000). Gender differences in processing information for making self-assessments of health. *Psychosomatic Medicine, 62*(3), 354-364.

Benyamini, Y., Leventhal, E. A., & Leventhal, H. (2003). Elderly people's ratings of the importance of health-related factors to their self-assessments of health. *Social Science & Medicine (1982), 56*(8), 1661-1667.

Bishop, G. D. (1991). *Understanding the understanding of illness: Lay disease representations. Mental representation in health and illness*. New York, NY: Springer.

Bjorner, J. B., Kristensen, T. S., Orth-Gomér, K., Tibblin, G., Sullivan, M., & Westerholm, P. (1996). *Self-rated health: A useful concept in research*. Prevention and Clinical Medicine. Swedish Council for Planning and Coordination of Research.

Bland, J. M., & Altman, D. G. (2000). Statistics notes. the odds ratio. *BMJ (Clinical Research Ed.), 320*(7247), 1468.

Blaxter, M. (1989). A comparison of measures of inequality in morbidity. *Health inequalities in European countries. Aldershot: Gower*, *199*, 230.

Blaxter, M. (1990). *Health and lifestyles*. London: Tavistock.

Blaxter, M. (2003). *Health and lifestyles.* Abingdon-on-Thames: Routledge.

Boccia, S. (2012). Personalized health care: The hope beyond the hype. *Italian Journal of Public Health, 9*(4).

Bohannon, R. W. (2001). Dynamometer measurements of hand-grip strength predict multiple outcomes. *Perceptual and Motor Skills, 93*(2), 323-328.

Borawski, E. A., Kinney, J. M., & Kahana, E. (1996). The meaning of older adults' health appraisals: Congruence with health status and determinant of mortality. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 51*(3), S157-S170.

Bosworth, H. B., Siegler, I. C., Brummett, B. H., Barefoot, J. C., Williams, R. B., Clapp-Channing, N. E., & Mark, D. B. (1999). The association between self-rated health and mortality in a well-characterized sample of coronary artery disease patients. *Medical Care*, 1226-1236.

Bunker, J. P., Gomby, D. S., & Kehrer, B. H. (1989). *Pathways to health: The role of social factors.* Menlo Park, CA: Henry J. Kaiser Family Foundation.

Burström, B., & Fredlund, P. (2001). Self rated health: Is it as good a predictor of subsequent mortality among adults in lower as well as in higher social classes? *Journal of Epidemiology & Community Health,* 836-840 5p.

Chan, Y. Y., Teh, C. H., Lim, K. K., Lim, K. H., Yeo, P. S., Kee, C. C., . . . Ahmad, N. A. (2015). Lifestyle, chronic diseases and self-rated health among malaysian adults: Results from the 2011 national health and morbidity survey (NHMS). *BMC Public Health, 15*(1), 1.

Chandola, T., & Jenkinson, C. (2000). Validating self-rated health in different ethnic groups. *Ethnicity & Health, 5*(2), 151-159 9p.

Christian, L. M., Glaser, R., Porter, K., Malarkey, W. B., Beversdorf, D., & Kiecolt-Glaser, J. K. (2011). Poorer self-rated health is associated with elevated inflammatory markers among older adults. *Psychoneuroendocrinology, 36*(10), 1495-1504.

Cockerham, W. C., Sharp, K., & Wilcox, J. A. (1983). Aging and perceived health status. *Journal of Gerontology, 38*(3), 349-355.

Collins, C. E., Young, A. F., & Hodge, A. (2008). Diet quality is associated with higher nutrient intake and self-rated health in mid-aged women. *Journal of the American College of Nutrition, 27*(1), 146-157.

COM(95). Communication from the commission concerning a community action programme on health monitoring in the context of the framework for action in the field of public health. 449, 16.10.1995 Cong. (1995).

Conti, G., Heckman, J., & Urzua, S. (2010). The education-health gradient. *The American Economic Review, 100*(2), 234-238.

Cooper, R., Kuh, D., & Hardy, R. (2010). Objectively measured physical capability levels and mortality: Systematic review and meta-analysis. *BMJ (Clinical Research Ed.), 341*, c4467-c4467.

Coste, S. C. (2006). Allostasis, homeostasis, and the costs of physiological adaptation. *Quarterly Review of Biology, 81*(1), 55-56.

Courtenay, W. H. (2000). Engendering health: A social constructionist examination of men's health beliefs and behaviors. *Psychology of Men & Masculinity, 1*(1), 4.

Cummings, S. R., Nevitt, M. C., Browner, W. S., Stone, K., Fox, K. M., Ensrud, K. E., . . . Vogt, T. M. (1995). Risk factors for hip fracture in white women. study of osteoporotic fractures research group. *The New England Journal of Medicine, 332*(12), 767-773.

Cutler, D. M., & Lleras-Muney, A. (2006). *Education and Health: Evaluating Theories and Evidence* (No. w12352). National Bureau of Economic Research.

Dawson, D., & Adams, P. (1987). *National center for health statistics: Current estimates from the national health interview survey, United States, 1986. Vital and health statistics.* (10 (164) ed.). Washington. U.S: Public Health Service.

De Bruin, A. (1996). *Health interview surveys: Towards international harmonization of methods and instruments. WHO regional publications, european series, no. 58.* ERIC.

De Looper, M., & Lafortune, G. (2009). Measuring disparities in health status and in access and use of health care in OECD countries. *OECD Health Working Papers*, (43), 1.

De Vet, H. C., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in medicine: A practical guide*. Cambridge: Cambridge University Press.

Delpierre, C., Datta, G. D., Kelly-Irving, M., Lauwers-Cances, V., Berkman, L., & Lang, T. (2012). What role does socio-economic position play in the link between functional limitations and self-rated health: France vs. USA? *European Journal of Public Health, 22*(3), 317-321.

Delpierre, C., Lauwers-Cances, V., Datta, G. D., Lang, T., & Berkman, L. (2009). Using self-rated health for analysing social inequalities in health: A risk for underestimating the gap between socioeconomic groups? *Journal of Epidemiology and Community Health, 63*(6), 426-432.

Derogatis, L. R., & Unger, R. (2010). Symptom checklist-90-revised. *Corsini Encyclopedia of Psychology.* Hoboken, NJ*:* Wiley.

DeSalvo, K. B., Bloser, N., Reynolds, K., He, J., & Muntner, P. (2006). Mortality prediction with a single general self-rated health question. *Journal of General Internal Medicine, 21*(3), 267-275.

Djerf, K., Laiho, J., Lehtonen, R., Härkänen, T., & Knekt, P. (2008). Weighting and statistical analysis. In Heistaro, S. (Ed.), *Methodology report: Health 2000 survey* (pp. 182). Helsinki: Kansanterveyslaitos.

Dowd, J. B., & Todd, M. (2011). Does self-reported health bias the measurement of health inequalities in U.S. adults? evidence using anchoring vignettes from the health and retirement study. *The Journals of Gerontology.Series B, Psychological Sciences and Social Sciences, 66*(4), 478-489.

Dowd, J. B., & Zajacova, A. (2007). Does the predictive power of self-rated health for subsequent mortality risk vary by socioeconomic status in the US? *International Journal of Epidemiology, 36*(6), 1214-1221.

Dowd, J. B., & Zajacova, A. (2010). Does self-rated health mean the same thing across socioeconomic groups? evidence from biomarker data. *Annals of Epidemiology, 20*(10), 743-749.

EHIS. (2010). EHIS wave 1 guidelines. Retrieved from http://ec.europa.eu/eurostat/documents/203647/203710/EHIS_wave_1_guidelines.pdf/ffbeb62c-8f64-4151-938c-9ef171d148e0

Era, P., Jokela, J., & Heikkinen, E. (1986). Reaction and movement times in men of different ages: A population study. *Perceptual and Motor Skills, 63*(1), 111-130.

Eriksson, I., Unden, A. L., & Elofsson, S. (2001). Self-rated health. comparisons between three different measures. results from a population study. *International Journal of Epidemiology, 30*(2), 326-333.

ESS. (2016). The european social survey. Retrieved from http://www.europeansocialsurvey.org/

Etilé, F., & Milcent, C. (2006). Income-related reporting heterogeneity in self-assessed health: Evidence from france. *Health Economics, 15*(9), 965-981.

Eurostat. (2016). Eurostat. your key to european statistics. Retrieved from http://ec.europa.eu/eurostat

Fairlie, R. W. (2005). An extension of the blinder-oaxaca decomposition technique to logit and probit models. *Journal of Economic and Social Measurement, 30*(4), 305-316.

Farmer, M. M., & Ferraro, K. F. (1997). Distress and perceived health: Mechanisms of health decline. *Journal of Health and Social Behavior,* 298-311.

Fayers, P. M. (2005). *Assessing quality of life in clinical trials: Methods and practice.* Oxford: Oxford University Press.

Ferraro, K. F., & Kelley-Moore, J. A. (2001). Self-rated health and mortality among black and white adults: Examining the dynamic evaluation thesis. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 56*(4), S195-205.

Ferraro, K. F. (1980). Self-ratings of health among the old and the old-old. *Journal of Health and Social Behavior, 21*(4), 377-383.

Finch, B. K., Hummer, R. A., Reindl, M., & Vega, W. A. (2002). Validity of self-rated health among latino(a)s. *American Journal of Epidemiology, 155*(8), 755-759.

Fylkesnes, K., & Førde, O., H. (1991). The Tromsø study: Predictors of self-evaluated health--has society adopted the expanded health concept? *Social Science & Medicine (1982), 32*(2), 141-146.

Fylkesnes, K., & Førde, ,O.,H. (1992). Determinants and dimensions involved in self-evaluation of health. *Social Science & Medicine (1982), 35*(3), 271-279.

Garson, G. (2012). *Logistic regression: Statnotes.* Raleigh, NC: North Carolina State University Public Administration.

Goldman, N., Glei, D. A., & Chang, M. (2004). The role of clinical risk factors in understanding self-rated health. *Annals of Epidemiology, 14*(1), 49-57.

Goldman, N. (2001). Social inequalities in health: Disentangling the underlying mechanisms. In M. Weinstein, A. I. Hermalin, M. A. Stoto, M. (Eds.), (pp. 118-139). New York, NY: New York Academy of Sciences.

Goldstein, M. S., Siegel, J. M., & Boyer, R. (1984). Predicting changes in perceived health status. *American Journal of Public Health, 74*(6), 611-614.

Graubard, B. I., & Korn, E. L. (1999). Predictive margins with survey data. *Biometrics, 55*(2), 652-659.

Groot, W. (2000). Adaptation and scale of reference bias in self-assessments of quality of life. *Journal of Health Economics, 19*(3), 403-420.

Gujarati, D. N. (2003). *Basic econometrics*. 4th. New York, NY: McGraw-Hill.

Halford, C., Anderzén, I., & Arnetz, B. (2003). Endocrine measures of stress and self-rated health: A longitudinal study. *Journal of Psychosomatic Research, 55*(4), 317-320.

Heidrich, J., Liese, A. D., Löwel, H., & Keil, U. (2002). Self-rated health and its relation to all-cause and cardiovascular mortality in southern Germany. Results from the MONICA Augsburg cohort study 1984–1995. *Annals of Epidemiology, 12*(5), 338-345.

Heistaro, S. (2008). *Methodology report: Health 2000 survey.* Helsinki: Kansanter-veyslaitos.

Heistaro, S., Vartiainen, E., & Puska, P. (1996). Trends in self-rated health in Finland 1972–1992. *Preventive Medicine, 25*(5), 625-632.

Heistaro, S., Jousilahti, P., Lahelma, E., Vartiainen, E., & Puska, P. (2001). Self-rated health and mortality: A long term prospective study in eastern Finland. *Journal of Epidemiology and Community Health, 55*(4), 227-232.

Helakorpi, S., Uutela, A., Prättälä, R., & Puska, P. (1999). *Suomalaisen aikuisväestön terveyskäyttäytyminen ja terveys, kevät 1999*. Helsinki: Kansanterveyslaitoksen Julkaisuja B: 19/1999.

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression.* Hoboken, NY: John Wiley & Sons.

Huber, M., Knottnerus, J. A., Green, L., van der Horst, H., Jadad, A. R., Kromhout, D., . . . Smid, H. (2011). How should we define health? *BMJ (Clinical Research Ed.), 343*, d4163.

Huisman, M., van Lenthe, F., & Mackenbach, J. (2007). The predictive ability of self-assessed health for mortality in different educational groups. *International Journal of Epidemiology, 36*(6), 1207-1213.

Huttunen, J. (2012). *Mitä terveys on.* Helsinki: Lääkärikirja Duodecim, 15, 2012.

Iburg, K. M., Salomon, J. A., Tandon, A., & Murray, C. J. (2001). *Cross-population comparability of self-reported and physician-assessed mobility levels: Evidence from the third national health and nutrition examination survey*. *Global Programme on Evidence for Health Policy Discussion Paper.* Geneva: World Health Organization.

Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior, 38*(1), 21-37.

Idler, E. L., & Kasl, S. V. (1995). Self-ratings of health: Do they also predict change in functional ability? *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 50*(6), S344-S353.

Idler, E. L., Russell, L. B., & Davis, D. (2000). Survival, functional limitations, and self-rated health in the NHANES I epidemiologic follow-up study, 1992. first national health and nutrition examination survey. *American Journal of Epidemiology, 152*(9), 874-883.

Idler, E. L., & Angel, R. J. (1990). Self-rated health and mortality in the NHANES-I epidemiologic follow-up study. *American Journal of Public Health, 80*(4), 446-452.

Idler, E. L., & Kasl, S. (1991). Health perceptions and survival: Do global evaluations of health status really predict mortality? *Journal of Gerontology, 46*(2), S55-65.

Idler, E. L. (1993). Age differences in self assessments of health: Age changes, cohort differences, or survivorship? *Journal of Gerontology, 48*(6), S289-S300.

Idler, E. L., Hudson, S. V., & Leventhal, H. (1999). The meanings of self-ratings of health. *Research on Aging, 21*(3), 458.

Idler, E., Leventhal, H., McLaughlin, J., & Leventhal, E. (2004). In sickness but not in health: Self-ratings, identity, and mortality. *Journal of Health and Social Behavior, 45*(3), 336-356.

Jann, B. (2008). *Fairlie: Stata module to generate nonlinear decomposition of binary outcome differentials.* Boston, MA: Boston College Department of Economics: Statistical Software Components.

Jarczok, M. N., Kleber, M. E., Koenig, J., Loerbroks, A., Herr, R. M., Hoffmann, K., . . . Thayer, J. F. (2015). Investigating the associations of self-rated health: Heart rate variability is more strongly associated than inflammatory and other frequently used biomarkers in a cross sectional occupational sample. *PloS One, 10*(2), e0117196.

Johnson, T. P., Stallones, L., Garrity, T. F., & Marx, M. B. (1990). Components of self-rated health among adults: Analysis of multiple data sources. *International Quarterly of Community Health Education, 11*(1), 29-41.

Jorgensen, P., Langhammer, A., Krokstad, S., & Forsmo, S. (2015). Diagnostic labelling influences self-rated health. A prospective cohort study: The HUNT study, norway. *Family Practice, 32*(5), 492-499.

Jürges, H. (2007). True health vs response styles: Exploring cross-country differences in self-reported health. *Health Economics, 16*(2), 163-178.

Jylhä, M., Guralnik, J. M., Ferrucci, L., Jokela, J., & Heikkinen, E. (1998). Is self-rated health comparable across cultures and genders? *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 53*(3), S144-S152.

Jylhä, M. (2009). What is self-rated health and why does it predict mortality? towards a unified conceptual model. *Social Science & Medicine, 69*(3), 307-316.

Jylhä, M. (2011). *Self-rated health and subjective survival probabilities as predictors of mortality*. International handbook of adult mortality (pp. 329-344). New York, NY: Springer.

Jylhä, M., Leskinen, E., Alanen, E., Leskinen, A. L., & Heikkinen, E. (1986). Self-rated health and associated factors among men of different ages. *Journal of Gerontology, 41*(6), 710-717.

Kaplan, G. A., Strawbridge, W. J., Camacho, T., & Cohen, R. D. (1993). Factors associated with change in physical functioning in the elderly. A six-year prospective study. *Journal of Aging and Health, 5*(1), 140-153.

Kaplan, G. A., & Camacho, T. (1983). Perceived health and mortality: A nine-year follow-up of the human population laboratory cohort. *American Journal of Epidemiology, 117*(3), 292-304.

Kaplan, G., & Baron-Epel, O. (2003). What lies behind the subjective evaluation of health status? *Social Science & Medicine (1982), 56*(8), 1669-1676.

Kaplan, R. M., Bush, J. W., & Berry, C. C. (1976). Health status: Types of validity and the index of well-being. *Health Services Research, 11*(4), 478-507.

Karisto, A. (1981). *Sosiaalilääketiede ja yhteiskunta: Katsaus suomalaiseen terveyden sosiaalisia eroja koskevaan tutkimustoimintaan autonomian ajalta 1930-luvulle.* Helsinki: Helsingin yliopisto, Sosiaalipolitiikan laitos.

Karisto, A. (1984). *Hyvinvointi ja sairauden ongelma*. Helsinki: Publications of the Social Insurance Institution M/46.

Kennedy, A., Rogers, A., & Gately, C. (2005). Assessing the introduction of the expert patients programme into the NHS: A realistic evaluation of recruitment to a national lay-led self-care initiative. *Primary Health Care Research & Development (Sage Publications, Ltd.), 6*(2), 137-148.

Kervinen, H. (2013). *Akuutti sepelvaltimo-oireyhtymä ja sydäninfarkti.* Helsinki: Lääkärin Käsikirja Duodecim.

Kessler, R. C., & Üstün, T. B. (2004). The world mental health (WMH) survey initiative version of the world health organization (WHO) composite international diagnostic interview (CIDI). *International Journal of Methods in Psychiatric Research, 13*(2), 93-121.

Koskinen, S., Knekt, P., Alha, P., Rinne, S., Virtala, E., Rissanen, H., & Aromaa, A. (2008). *Data obtained from registers and register follow-up*. In S. Heistaro (Ed.), *Methodology report, health 2000 survey*. Helsinki: Publications of the National Public Health Institute B26.

Krause, N. M., & Jay, G. M. (1994). What do global self-rated health items measure? *Medical Care, 32*(9), 930-942.

Laiho, J., Djerf, K., & Lehtonen, R. (2008). *Sampling design*. In S. Heistaro (Ed.), *Methodology report, health 2000 survey.* Helsinki.: Publications of the National Public Health Institute B26.

Lallukka, T., Rahkonen, O., Lahelma, E., & Laaksonen, M. (2011). Sosiaaliluokkien välisten terveyserojen selittäminen ja niiden vähentämisen haaste. esimerkkinä työkyvyttömyyseläkkeelle joutuminen. *Janus Sosiaalipolitiikan Ja Sosiaalityön Tutkimuksen Aikakauslehti, 19*(4), 358-368.

Leinonen, R., Heikkinen, E., & Jylhä, M. (2001). Predictors of decline in self-assessments of health among older people--a 5-year longitudinal study. *Social Science & Medicine (1982), 52*(9), 1329-1341.

Leventhal, H., Diefenbach, M., & Leventhal, E. A. (1992). Illness cognition: Using common sense to understand treatment adherence and affect cognition interactions. *Cognitive Therapy and Research, 16*(2), 143-163.

Levkoff, S. E., Cleary, P. D., & Wetle, T. (1987). Differences in the appraisal of health between aged and middle-aged adults. *Journal of Gerontology, 42*(1), 114-120.

Lindeboom, M., & Van Doorslaer, E. (2004). Cut-point shift and index shift in self-reported health. *Journal of Health Economics, 23*(6), 1083-1099.

Lommi, J. (2013). *Sydämen krooninen vajaatoiminta*. Helsinki: Lääkärin Käsikirja Duodecim 24.

Lund, T., & Borg, V. (1999). Work environment and self-rated health as predictors of remaining in work 5 years later among danish employees 35–59 years of age. *Experimental aging research, 25(4), 429-434.*

Lundberg, O., & Manderbacka, K. (1996). Assessing reliability of a measure of self-rated health. *Scandinavian Journal of Social Medicine, 24*(3), 218-224.

Mackenbach, J. P. (2006). *Health inequalities: Europe in profile* Produced by COI for the Department of Health.

Mackenbach, J. P., Stirbu, I., Roskam, A. R., Schaap, M. M., Menvielle, G., Leinsalu, M., & Kunst, A. E. (2008). Socioeconomic inequalities in health in 22 european countries. *New England Journal of Medicine, 358*(23), 2468-2481.

Mackenbach, J. P., van den Bos, J., Joung, I. M., van de Mheen, H., & Stronks, K. (1994). The determinants of excellent health: Different from the determinants of ill-health? *International Journal of Epidemiology, 23*(6), 1273-1281.

Maddox, G. L. (1962). Some correlates of differences in self-assessment of health status among the elderly. *Journal of Gerontology.*

Maddox, G. L., & Douglass, E. B. (1973). Self-assessment of health: A longitudinal study of elderly subjects. *Journal of Health and Social Behavior, 14*(1), 87-93.

Manderbacka, K., Lundberg, O., & Martikainen, P. (1999). Do risk factors and health behaviours contribute to self-ratings of health? *Social Science & Medicine, 48*(12), 1713-1720.

Manderbacka, K. (1998). Examining what self-rated health question is understood to mean by respondents. *Scandinavian Journal of Social Medicine, 26*(2), 145-153.

Manderbacka, K., Lahelma, E., & Martikainen, P. (1998). Examining the continuity of self-rated health. *International Journal of Epidemiology, 27*(2), 208-213.

Manderbacka, K. (1998a). *Questions on survey questions on health*. Stockholm: Swedish Institute for Social Research, 1998.

Manderbacka, K. (1998b). How do respondents understand survey questions on ill-health? *European Journal of Public Health, 8*(4), 319.

Manor, O., Matthews, S., & Power, C. (2000). Dichotomous or categorical response? analysing self-rated health and lifetime social class. *International Journal of Epidemiology, 29*(1), 149-157.

Mäntyselkä, P. (2004). Krooninen kipu ja koettu terveys suomalaisilla. *Kipuviesti: Suomen Kivuntutkimusyhdistyksen Jäsenlehti,* (1), 10-11.

Martikainen, P., Lahelma, E., Marmot, M., Sekine, M., Nishi, N., & Kagamimori, S. (2004). A comparison of socioeconomic differences in physical functioning and perceived health among male and female employees in Britain, Finland and Japan. *Social Science & Medicine, 59*(6), 1287-1295.

Mathers, C., & Douglas, R. (1998). *Measuring progress in population health and well-being*. In R. Eckersley (Ed.), *Measuring progress: Is life getting better?* (pp. 125-155). Clayton, VIC: CSIRO Publishing.

McCullagh, P. (1980). Regression models for ordinal data. *Journal of the Royal Statistical Society.Series B (Methodological),* 109-142.

McDowell, I. (2006). *Measuring health: A guide to rating scales and questionnaires*. Oxford: Oxford university press.

Menard, S. (2009). *Logistic regression: From introductory to advanced concepts and applications.* Thousand Oaks*, CA:* Sage Publications.

Merrian-Webster Dictionary. (2016). "Health". Retrieved from http://www.merriam-webster.com/dictionary/health

Ministry of Social Affairs and Health. (2001). *Terveys 2015.* (No. 8). Helsinki: Brochures of the Ministry of Social Affairs and Health.

Mossey, J. M., & Shapiro, E. (1982). Self-rated health: A predictor of mortality among the elderly. *American Journal of Public Health, 72*(8), 800-808.

Moum, T. (1992). Self-assessed health among norwegian adults. *Social Science & Medicine, 35*(7), 935-947.

Myers, G. C. (1974). Differential mortality in the united states: A study in socioeconomic epidemiology. *American Journal of Sociology, 80*(2), 532-534.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity, 41*(5), 673-690.

Onadja, Y., Bignami, S., Rossier, C., & Zunzunegui, M. (2013). The components of self-rated health among adults in ouagadougou, burkina faso. *Population Health Metrics, 11*(1), 1.

Palosuo, H., Koskinen, S., Lahelma, E., Kostiainen, E., Prättälä, R., Martelin, T., . . . (Eds.). (2009). *Health inequalities in Finland. trends in socioeconomic health differences 1980-2005.* (9th ed.). Helsinki: Ministry of Social Affairs and Health Publications.

Piukka, K. (2008). *Piukka, K (2008) hypertension kriteeri terveys 2000 –tutkimuksessa.* (No. 02/2008). Helsinki: Kansanterveyslaitos.

Pope, G. C. (1988). Medical conditions, health status, and health services utilization. *Health Services Research, 22*(6), 857-877.

Prättälä, R., Koskinen, S., Martelin, T., Lahelma, E., Sihto, M., & Palosuo, H. (2007). Terveyserot ja niiden kaventamisen haaste. In H. Palosuo, S. Koskinen, E. Lahelma, R. Prättälä, T. Martelin, A. Ostamo, . . . E. Linnanmäki (Eds.), *Terveyden eriarvoisuus suomessa sosioekonomisten terveyserojen muutokset 1980–2005.* (pp. 15-24). Helsinki: Ministry of Social Affairs and Health Publications.

Rahkonen, O., Talala, K., Laaksonen, M., Lahelma, E., Prättälä, R., & Uutela, A. (2004). Suomalaisten koettu terveys parantunut, terveyden koulutuserot säilyneet 1979-2002. *Suomen Lääkärilehti - Finlands Läkartidning, 59*(20), 2159-2163.

Rantanen, T., Guralnik, J. M., Foley, D., Masaki, K., Leveille, S., Curb, J. D., & White, L. (1999). Midlife hand grip strength as a predictor of old age disability. *JAMA, 281*(6), 558-560.

Regidor, E., Guallar-Castillon, P., Gutierrez-Fisac, J. L., Banegas, J. R., & Rodriguez-Artalejo, F. (2010). Socioeconomic variation in the magnitude of the association between self-rated health and mortality. *Annals of Epidemiology, 20*(5), 395-400.

Reunanen, A. (2005). *Verenkiertoelinten sairaudet*. *Suomalaisten Terveys.* Helsinki: Duodecim, Stakes, Kansanterveyslaitos.

Rijk, J. M., Roos, P. R., Deckx, L., van, d. A., & Buntinx, F. (2015). Prognostic value of handgrip strength in people aged 60 years and older: A systematic review and meta-analysis. *Geriatrics & Gerontology International, 16(1), 5-20.*

Rodin, J., & McAvay, G. (1992). Determinants of change in perceived health in a longitudinal study of older adults. *Journal of Gerontology, 47*(6), P373-84.

Roine, R. (2013). *Aivoinfarkti*. Helsinki: Duodecim Lääkärin Käsikirja.

Rowan, K. (1994). Global questions and scores. *Measuring Health and Medical Outcomes,* 54-76.

Sainio, P., Aromaa, A. & Koskinen, S. (2008). Terveys 2000 viitearvot. Retrieved from http://www.terveys2000.fi/viitearvot/viitearvo.html

Salomon, J. A., Tandon, A., & Murray, C. J. L. (2004). Comparability of self rated health: Cross sectional multi-country survey using anchoring vignettes. *BMJ (Clinical Research Ed.), 328*(7434), 258-258.

Schnittker, J. (2005). When mental health becomes health: Age and the shifting meaning of Self-Evaluations of general health. *Milbank Quarterly, 83*(3), 397-423.

Schnittker, J. (2009). Mirage of health in the era of biomedicalization: Evaluating change in the threshold of illness, 1972–1996. *Social Forces, 87*(4), 2155-2182.

Schultz-Larsen, K., Avlund, K., & Kreiner, S. (1992). Functional ability of community dwelling elderly. criterion-related validity of a new measure of functional ability. *Journal of Clinical Epidemiology, 45*(11), 1315-1326.

Sen, A. (2002). Perception versus observation. *British Medical Journal, 324*, 859-860.

Shadbolt, B. (1997). Some correlates of self-rated health for australian women. *American Journal of Public Health, 87*(6), 951-956.

Shadbolt, B., Barresi, J., & Craft, P. (2002). Self-rated health as a predictor of survival among patients with advanced cancer. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology, 20*(10), 2514-2519.

SHARE. (2016). The survey of health, ageing and retirement in europe (SHARE). Retrieved from http://www.share-project.org/

Shavers, V. L. (2007). Measurement of socioeconomic status in health disparities research. *Journal of the National Medical Association, 99*(9), 1013-1023.

Shmueli, A. (2003). Socio-economic and demographic variation in health and in its measures: The issue of reporting heterogeneity. *Social Science & Medicine, 57*(1), 125-134.

Singh-Manoux, A., & Marmot, M. (2005). Role of socialization in explaining social inequalities in health. *Social Science & Medicine (1982), 60*(9), 2129-2133.

Singh-Manoux, A., Gueguen, A., Martikainen, P., Ferrie, J., Marmot, M., & Shipley, M. (2007). Self-rated health and mortality: Short- and long-term associations in the whitehall II study. *Psychosomatic Medicine, 69*(2), 138-143.

Sinning, M., Hahn, M., & Bauer, T. K. (2008). The blinder-oaxaca decomposition for nonlinear regression models. *The Stata Journal, 8*(4), 480-492.

Smith, A. M., Shelley, J. M., & Dennerstein, L. (1994). Self-rated health: Biological continuum or social discontinuity? *Social Science & Medicine (1982), 39*(1), 77-83.

Smith, P. M., Glazier, R. H., & Sibley, L. M. (2010). The predictors of self-rated health and the relationship between self-rated health and health service needs are similar across socioeconomic groups in canada. *Journal of Clinical Epidemiology, 63*(4), 412-421.

Sohn, H. (2015). Education and self-awareness of health: Toward a better understanding of self-rated health. *Education, Social Factors, and Health Beliefs in Health and Health Care Services (Research in the Sociology of Health Care, Volume 33) Emerald Group Publishing Limited, 33*, 85-103.

Stata Manual. (2016). Logit. Retrieved from http://www.stata.com/manuals13/rlogit.pdf

Stenholm, S., Kivimäki, M., Jylhä, M., Kawachi, I., Westerlund, H., Pentti, J., . . . Vahtera, J. (2016). Trajectories of self-rated health in the last 15 years of life by cause of death. *European Journal of Epidemiology, 31*(2), 177-185.

Steyerberg, E. (2008). *Clinical prediction models: A practical approach to development, validation, and updating*. New York, NY: Springer Science & Business Media.

Terveyttä kaikille vuoteen 2000. Suomen terveyspolitiikan pitkän aikavälin tavoite- ja toimintaohjelma. Helsinki: Valtion painatuskeskus (1986).

Sosiaali- ja terveysministeriö (STM). valtioneuvoston periaatepäätös terveys 2015 -kansanterveysohjelmasta. Helsinki: Sosiaali- ja terveysministeriön julkaisuja 4, Helsinki.Cong. (2001).

Subramanian, S. V., Huijts, T., & Avendano, M. (2010). Self-reported health assessments in the 2002 world health survey: How do they correlate with education? *Bulletin of the World Health Organization, 88*(2), 131-138.

Suchman, E. A., Phillips, B. S., & Streib, G. F. (1958). An analysis of the validity of health questionnaires. *Social Forces, 36*(3), 223-232.

Suls, J., Marco, C. A., & Tobin, S. (1991). The role of temporal comparison, social comparison, and direct appraisal in the elderly's Self-Evaluations of health. *Journal of Applied Social Psychology, 21*(14), 1125-1144.

Szwarcwald, C. L., Damacena, G. N., Souza Júnior, Paulo Roberto Borges de, Almeida, Wanessa da Silva de, Lima, Lilandra Torquato Medrado de, Malta, D. C., . . . Pereira, C. A. (2015). Determinants of self-rated health and the influence of healthy behaviors: Results from the national health survey, 2013. *Revista Brasileira De Epidemiologia, 18*, 33-44.

Tarkiainen, L., Martikainen, P., Laaksonen, M., & Valkonen, T. (2012). Trends in life expectancy by income from 1988 to 2007: Decomposition by age and cause of death. *Journal of Epidemiology and Community Health, 66*(7), 573-578.

Terveyden ja hyvinvoinnin laitos. (2016). Hyvinvointikompassi. Hyvinvoinnin indikaattorit. Retrieved from http://hyvinvointikompassi.thl.fi/fi/web/hyvinvointikompassi/hyvinvointi-profiili;jsessionid=4B4774BDD144D8E511ADFDBA91228FC4

Thomas, D., & Frankenberg, E. (2002a). Health, nutrition and prosperity: A microeconomic perspective. *Bulletin of the World Health Organization, 80*(2), 106-113.

Thomas, D., & Frankenberg, E. (2002b). The measurement and interpretation of health in social surveys. *Measurement of the Global Burden of Disease,* 387-420.

Tibblin, G., Cato, K., & Svardsudd, K. (1990). Goteborg quality of life study of men born in 1913 and 1923--age, sex, job satisfaction and cardiovascular diseases. *Scandinavian Journal of Primary Health Care.Supplement, 1*, 39-45.

Tjur, T. (2009). Coefficients of determination in logistic regression models—A new proposal: The coefficient of discrimination. *The American Statistician, 63*(4), 366-372.

Tolonen, H., Kuulasmaa, K., Laatikainen, T., & Wolf, H. (2002). *Recommendation for indicators, international collaboration, protocol and manual of operations for chronic disease risk factor surveys. Project EHRM, European Health Risk Monitoring Project.* Helsinki, Finland: National Public Health Institute.

Townsend, P., & Davidson, N. (1982). Inequality in health: The black report.

Van Doorslaer, E., & Gerdtham, U. (2003). Does inequality in self-assessed health predict inequality in survival by income? evidence from swedish data. *Social Science & Medicine, 57*(9), 1621-1629.

Verschuuren, M., Gissler, M., Kilpeläinen, K., Tuomi-Nikula, A., Sihvonen, A., Thelen, J., . . . Prochorskas, R. (2013). Public health indicators for the EU: The joint action for ECHIM (european community health indicators & monitoring). *Archives of Public Health, 71*(1), 1.

Viitasalo, J., Era, P., Leskinen, A., & Heikkinen, E. (1985). Muscular strength profiles and anthropometry in random samples of men aged 31–35, 51–55 and 71–75 years. *Ergonomics, 28*(11), 1563-1574.

Von Wright, G. (1963). *The varieties of goodness.* London: Thoemmes Press.

Vuorisalmi, M., Lintonen, T., & Jylhä, M. (2005). Global self-rated health data from a longitudinal study predicted mortality better than comparative self-rated health in old age. *Journal of Clinical Epidemiology, 58*(7), 680-687.

Vuorisalmi, M., Lintonen, T., & Jylhä, M. (2006). Comparative vs global self-rated health: Associations with age and functional ability. *Aging Clinical and Experimental Research, 18*(3), 211-217.

Vuorisalmi, M., Pietilä, I., Pohjolainen, P., & Jylhä, M. (2008). Comparison of self-rated health in older people of st. petersburg, russia, and tampere, finland: How sensitive is SRH to cross-cultural factors? *European Journal of Ageing, 5*(4), 327-334.

Ware Jr, J. E., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36): I. conceptual framework and item selection. *Medical Care,* 473-483.

Ware, J. E. (1986). The assessment of health status. *Applications of Social Science to Clinical Medicine and Health Policy,* 204-228.

Warnoff, C., Lekander, M., Hemmingsson, T., Sorjonen, K., Melin, B., & Andreasson, A. (2016). Is poor self-rated health associated with low-grade inflammation in 43 110 late adolescent men of the general population? A cross-sectional study. *BMJ Open, 6*(4), e009440.

Weinberger, M., Darnell, J. C., Tierney, W. M., Martz, B. L., Hiner, S. L., Barker, J., & Neill, P. J. (1986). Self-rated health as a predictor of hospital admission and nursing home placement in elderly public housing tenants. *American Journal of Public Health, 76*(4), 457-459.

Willems, S., De Maesschalck, S., Deveugele, M., Derese, A., & De Maeseneer, J. (2005). Socio-economic status of the patient and doctor–patient communication: Does it make a difference? *Patient Education and Counseling, 56*(2), 139-146.

Wolinsky, F. D., Callahan, C. M., Fitzgerald, J. F., & Johnson, R. J. (1993). Changes in functional status and the risks of subsequent nursing home placement and death. *Journal of Gerontology, 48*(3), S94-101.

Woolfolk, R. L., & Allen, L. A. (2007). *Treating somatization: A cognitive-behavioral approach*. New York, NY: Guilford Press.

World Health Organization. (1948). WHO definition of health. Retrieved from: http://www.who.int/about/definition/en/print.html

World Health Organization. (1996). *Statistics Netherlands. Health interview surveys: Towards international harmonization of methods and instruments.* (WHO Regional Publications No. 58). Copenhagen: WHO Regional Office for Europe.

World Health Organization. (2016). BMI classification. Retrieved from http://apps.who.int/bmi/index.jsp?introPage=intro_3.html

Wu, S., Wang, R., Zhao, Y., Ma, X., Wu, M., Yan, X., & He, J. (2013). The relationship between self-rated health and objective health status: A population-based study. *BMC Public Health, 13*(1), 1.

Xue, Q., Walston, J. D., Fried, L. P., & Beamer, B. A. (2011). Prediction of risk of falling, physical disability, and frailty by rate of decline in grip strength: The women's health and aging study. *Archives of Internal Medicine, 171*(12), 1119-1121.

Yki-Järvinen, H., & Tuomi, T. (2013). *Diabeteksen määritelmä, erotusdiangnoosi ja luokitus.* Helsinki: Lääkärin käsikirja Duodecim.

Zborowski, M. (1952). Cultural components in responses to Pain1. *Journal of Social Issues, 8*(4), 16-30.

Zijlema, W. L., Stolk, R. P., Löwe, B., Rief, W., White, P. D., & Rosmalen, J. G. (2013). How to assess common somatic symptoms in large-scale studies: A systematic review of questionnaires. *Journal of Psychosomatic Research, 74*(6), 459-468.

Zola, I. (1966). Culture and symptoms: An analysis of patients' presenting complaints. *American Sociological Review, 31*(5), 615-630.

Zola, I. K. (1973). Pathways to the doctor: From person to patient. *Social Science & Medicine, 7*(9), 677-689.

Appendix 1. Abbreviations

BMI Body mass index (kg/m²)

CI Confidence interval

CVD Cardiovascular diseases

GHQ-12 12-item General Health Questionnaire

HDL High-density lipoprotein cholesterol

LDL Low-density lipoprotein cholesterol

OR Odds ratio

SAH Self-assessed health

SES Socio-economic status

SRH Self-rated health

SRQ20 Self-reporting questionnaire

Appendix 2. Literature search strategy.

The main target of the literature search strategy was global self-rated health measured using a 5-point Likert scale. Omitted were studies that address: a) comparative self-rated health, b) the association between self-rated health and some variable unrelated to physical or mental health such as self-assessed happiness or issues of causality related to income and self-rated health, and c) studies that target patient groups only. Studies that targeted ethnic, socioeconomic or age groups were included. Included were results in the following languages: Finnish, English, Swedish, French, Spanish and Italian.

The included databases were Ebsco databases (joint searches covered MedLine, AgeLine, PsycInfo, SocIndex with full text, Cinahl), Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, Medline (OVID), OT-seeker, Medic, Arto, Melinda and Theseus. The preliminary number of identified references was 354 and the final number after removing duplicates 192.

Example search strategies:
CINAHL, SocINDEX with Full Text, AgeLine, PsycINFO, MEDLINE (joint search) on 15.01.2016. TI "self-rated health" AND TI (psychometric* or clinimetr* or inter-rater or validity or validat* or reliability or reproducib* or sensitivity or specificity or feasible or feasibility or "minimal detectable change" or "significant change" or responsiveness or "predictive value of tests" or "observer variation" or “accuracy” or predict* or reporting heterogeneity or reporting style) AND (“education” or “socio-economic status” or “reporting heterogeneity”).

Web of science on 12.02.2016. TOPIC: (("self-rated health" or “self-assessed health”) AND (psychometric* or clinimetr* or inter-rater or validity or validati* or reliability or reproducib* or sensitivity or specificity or feasible or feasibility or "minimal detectable change" or "clinically significant change" or responsiveness or "predictive value of tests" or "observer variation") AND (education or socio-economic status or reporting heterogeneity)

Appendix 3. Clinical health variables

The second health category is clinical health: the existence of any documented clinical conditions and objective biological risk factors. The final dichotomized items that indicate the existence of a specific clinical condition was formed using disease diagnosis combination items formed by the Health 2000 study team. The existence of a disease was identified by the Health 2000 team using several sources: the home interview, the symptom interview, the medical health examination section and relevant registers such as the HILMO-system and medication registers from 1964 to 2004. Based on the previously mentioned, specific diseases were recoded, often as some variation of the following: “no diagnosis”, “possible diagnosis”, “certain diagnosis” and “patient has received medication for said disease after 2000/2001 as registered in the HILMO-system”.

For instance, in the case of heart failure the final, dichotomised disease item was formed as follows. When the respondent had not been registered as suffering from heart failure in the Hilmo-system, had not received special medical reimbursements for the treatment of heart failure prior to 2001 and received neither probable, nor plausible diagnosis in the medical examination, the respondent was recoded as “No heart failure”.

”Certain diagnosis” comprised of respondents who had been diagnosed with the disease in the medical examination and had some other source (self-report, received medical reimbursements, hospitalised for said condition or treatment otherwise recorded in the Hilmo-system) support the diagnosis.

The “possible diagnosis” –type was rare. With the selected disease items, it was assigned to less than 2% of the respondents. The respondent was usually recoded as a possible case if he or she had been treated in a hospital for the disease in question prior to the examination but the clinician’s assessment did not absolutely confirm the diagnosis, or vice versa.

A third option existed for some of the disease items. In the case of heart failure, it was formulated as: “Some indication of possible heart failure, unsuccessful to secure certain or plausible diagnosis”. This code was assigned to respondents with some singular indicators of the disease but a negative diagnosis by the medical examiner or when no prior treatment had been received for the condition. In most of these cases, the respondents subjective view had been that he or she had received a prior diagnose for the disease; otherwise the case would have been coded as a negative. In some cases, the study participant had received reimbursements for medication for the disease but the field clinician did not concur. This category is small. For heart failure, 0,47 % of the respondents (38 individuals) were included in this category.

All the disease items also included a category for those respondents who were diagnosed with the disease in question after the Health 2000 medical inspection as indicated by the HILMO medical care registration system or other register data sources. The data does not include information on the specific date of the diagnosis.

Appendix 4. Indicator correlation matrix.

Table 7. Correlation matrix, health indicators.

	Poor SRH	Smoking	Activity	Vegetab.	Obesity	Cholester.	Hypertens.	Musculoskel.	Cardiovasc.	Diabetes	Subj. func.	Tests func.	Clinician's	Somatist.	Disoders
Poor SRH	1.00														
Daily smoker	0.04**	1.00													
Low physical activity	0.05**	0.15**	1.00												
Low vegetable consumption	0.07**	0.12**	0.12**	1.00											
Obesity	0.10**	-0.04**	0.04**	0.04**	1.00										
High cholesterol	0.04**	0.03*	0.05**	0.06**	0.15**	1.00									
Hypertension	0.12**	-0.09**	0.02	0.06**	0.23**	0.15**	1.00								
Musculoskeletal disorders	0.23**	-0.02	-0.00	0.03*	0.11**	0.07**	0.14**	1.00							
Cardiovascular diseases	0.24**	-0.07**	-0.01	0.04**	0.08**	0.03*	0.19**	0.15**	1.00						
Diabetes	0.13**	-0.02	0.04**	0.01	0.16**	0.05**	0.17**	0.07**	0.14**	1.00					
Subjective troubles	**0.31****	0.02	0.06**	0.10**	0.20**	0.03**	0.14**	0.21**	0.16**	0.12**	1.00				
Low ability test score	0.14**	0.08**	0.03	0.05**	0.03	-0.01	-0.04**	0.06**	-0.00	0.07**	0.17**	1.00			
Clinician: limitations	**0.36****	-0.03*	0.06**	0.08**	0.16**	0.06**	0.20**	**0.41****	**0.33****	0.16**	**0.32****	0.10**	1.00		
High somatisation	**0.43****	0.01	0.03*	0.07**	0.12**	0.03*	0.12**	0.24**	0.21**	0.09**	**0.32****	0.12**	**0.33****	1.00	
Diagnosed disorder	0.13**	0.10**	0.04**	0.04**	0.01	-0.03*	-0.02	0.02	-0.05**	-0.01	0.07**	0.05**	0.02	0.11**	1.00

** p<,01 * p<,05