Salla Savela

PHYSICAL ACTIVITY IN MIDLIFE AND HEALTH-RELATED QUALITY OF LIFE, FRAILTY, TELOMERE LENGTH AND MORTALITY IN OLD AGE
SALLA SAVELA

PHYSICAL ACTIVITY IN MIDLIFE AND HEALTH-RELATED QUALITY OF LIFE, FRAILTY, TELOMERE LENGTH AND MORTALITY IN OLD AGE

Academic dissertation to be presented with the assent of the Doctoral Training Committee of Health and Biosciences of the University of Oulu for public defence in Auditorium F202 of the Department of Pharmacology and Toxicology (Aapistie 5 A), on 13 December 2014, at 12 noon

UNIVERSITY OF OULU, OULU 2014
Physical inactivity is an increasing health problem worldwide and one of the most important risk factors for global premature mortality. Along with gained years, it is highly essential for both individuals and society that elderly persons perceive their health as good and are able to cope well with everyday life and enjoy it. Hence, our aim was to investigate the long-term associations of midlife physical activity (PA) with health-related quality of life (HRQL), frailty stage, leukocyte telomere length (LTL), a potential indicator of biological ageing, and mortality.

The Helsinki Businessmen Study cohort has been followed since the 1960s and it has been studied from several perspectives over the years. In the baseline examinations in 1974, 782 healthy men (mean age 48) completed a questionnaire about their PA pattern. These men form the population of this study. According to their global description of their PA pattern, the men were categorized into low, moderate and high PA groups.

After a 26-year follow-up in 2000, the HRQL and disease prevalences were appraised in 552 (mean age 73) men using a postal questionnaire including the RAND-36 instrument. PA was significantly associated with better physical function, one of the eight domains of HRQL. From the same questionnaire, frailty stage was appraised for 514 men using modified Fried’s criteria: weight loss, physical inactivity, low vitality and physical weakness. After adjusting for CVD risk factors at baseline, the risk for frailty was 80% lower in the high PA group compared with the low PA group.

After a 29-year follow-up in 2003, 204 randomly selected survivors were invited to laboratory tests. The mean LTL was longer in the moderate PA group than in the low PA group and, contrary to our hypothesis, in the high PA group.

We found that midlife PA was related to mortality during a 34-year follow-up and the protective effect of PA was independent of the cardiovascular disease (CVD) risk factors, including body mass index, age, cholesterol, glucose, systolic blood pressure and smoking at baseline.

In conclusion, the overall results of this study lend support to the view that those who have adopted a physically active lifestyle in middle age have better physical function and lower risk of frailty in old age and lower mortality.

Keywords: exercise, frailty, mortality, physical activity, quality of life, social class, telomere
Savela, Salla, Keski-iän liikunnan yhteys terveyteen liittyvään elämänlaatuun, hauraus-raihnassa -oireyhtymään, telomeerien pituuteen sekä kuolleisuuteen vanhalla iällä.

Oulun yliopiston tutkijakoulu; Oulun yliopisto, Lääketieteellinen tiedekunta, Terveystieteiden laitos, Geriatrics; Oulun yliopistollinen sairaala; Helsingin yliopisto, Lääketieteellinen tiedekunta, Geriatrian klinikka; Helsingin yliopistollinen keskussairala

*Acta Univ. Oul. D 1276, 2014*

Oulun yliopisto, PL 8000, 90014 Oulun yliopisto

**Tiivistelmä**


Vuonna 1974, Helsingin Johtajat -tutkimuksen alkuvaiheessa, 782 terveitä miestä (keski-ikä 48 vuotta) täytti liikuntatottumuksiaan koskevan kyselylomakkeen. He muodostavat tämän tutkimuksen aineiston. Liikuntaharrastuksestaan an taman yleiskuvauksen mukaan miehet jaettiin matalan, keskitason ja korkean liikuntatason ryhmiin.


**Asiasanat:** elämänlaatu, fyysinen aktiivisuus, hauraus-raihnassa -oireyhtymä, kuolleisuus, sosiaaliluokka, telomeeri
To my family and friends
Acknowledgements

This doctoral research study was carried out at the University of Oulu, Institute of Health Sciences, during the years 2009-2014. The study is based on data of the Helsinki Businessmen Study, a still ongoing cohort study for which data were gathered in 1960-1970.

I would like to express my sincere thanks to all of those who have, in one way or another, taken part in this project. Particularly, I would like to thank the following persons.

My deepest gratitude goes to my principal supervisor, Professor Timo Strandberg, at the University of Helsinki, Helsinki University Central Hospital and the University of Oulu, who introduced me to the subject and framework of the Helsinki Businessmen Study. His constant support, advice and scientific guidance throughout these years has been invaluable. I feel great respect for his broad knowledge of geriatrics and medicine as well as his long-term research work and expertise in statistics. I am also grateful for his guidance to all of us who are specialized in geriatrics.

I would like to express my sincere gratitude to my other supervisor, Docent Pentti Koistinen at the Oulu City Hospital, for his guidance and encouragement from the very beginning of this work to the comments on this thesis. He introduced me to the clinical geriatric rehabilitation at the Oulu City Hospital, and his vast knowledge and experience in that area as well as sports medicine has been valuable in this research project.

My sincere thanks go also to my co-authors Professors Reijo Tilvis, Kaisu Pitkälä, Veikko Salomaa and Frej Fyhrquist, Ph.D. Arto Strandberg, Docent Outi Saijonmaa, Ph.D. Satu Stenholm and the late Professor Tatu Miettinen. Their contribution and comments to the articles have been extremely useful and instructive. Many of them have been working for years with the Helsinki Businessmen Study cohort and I am grateful for the excellent research they have done and the opportunity to work with them. I give special thanks to Professor Frej Fyhrquist, who carried out the telomere length measurements.

All the men who have taken part in the Helsinki Businessmen Study are acknowledged for making this study possible. Similarly, I would like to thank all the workers at the Institute of Occupational Health in Helsinki and Helsinki University Hospital who have contributed to the Helsinki Businessmen Study from the very beginning up to the present. The staff of Oulu University Graduate
School, Institute of Health Sciences and Faculty of Medicine are acknowledged for friendly and helpful collaboration.

Docents Mikaela von Bonsdorff and Jouko Laurila, the official reviewers of this thesis, are acknowledged for their constructive criticism and advice. Their comments really helped me to improve the final form of thesis. Similarly I wish to thank MA Anna Vuolteenaho, for the extensive and careful revision of the English of this thesis.

I have been fortunate to work also in a clinical setting during this project. All my colleagues at the Oulu City Hospital are warmly acknowledged. It is a pleasure to work with you! I want to give special thanks to my superiors Kristiina Pussinen, Anneli Vesa-Kurtti, Arja Mursu and Liisa Kylmänen for making it possible to finish this project alongside my clinical work and to Riitta Leskinen for practical advice concerning doctoral studies. Similarly, I would like to thank Professor Riitta Antikainen for an inspiring and encouraging ambience. I truly appreciate her will and capability to ensure that all the latest global trends in geriatrics are known at the Oulu City Hospital.

My warmest thanks are due to my mother and father, Eila and Seppo Niskanen, for their continuous love, support and encouragement in all phases of my life. They have helped me in various ways during this project and I owe particular thanks to my mother for so lovingly taking care of my children when needed. I want to thank my dear siblings and their families for all the unforgettable times we have spent together. The humor of the Niskanens always saves the day. My mother- and father-in-law and their family are acknowledged for their support and friendship.

My friends near and far deserve warm thanks for all the encouragement they have given me and the great moments we have shared. Just knowing that you are there gives me strength. I want to give special thanks to Raili, Karolina, Riikka-Maria, Leila, Elina, Kaisa and Päivi for offering me practical advice and help, opening their hearts and homes to my children, as well as taking me to various activities even during the most hectic phases of this project.

Finally, I wish to express my dearest thanks to my husband, Jyrki. Thank you for all your patience, support and endless love throughout these years. Your inborn interest and enthusiasm in my research topic, physical activity, has encouraged me to continue this project until this day. With love I thank our wonderful children, Benjamin, Lotte, Silvia and Louna, for filling my everyday life with joy and tenderness. Every time I think of you I remember what is truly important in life.
This work was financially supported by the Oulu University Hospital, the City of Oulu (EVO-funding), the Päivikki and Sakari Sohlberg Foundation, Konung Gustaf V:s och Drottning Victorias Frimurarestiftelse, the Uulo Arhio Foundation and the Finnish Medical Society Duodecim, which I gratefully acknowledge.

Oulu, November 2014

Salla Savela
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI</td>
<td>body mass index</td>
</tr>
<tr>
<td>BP</td>
<td>base pairs</td>
</tr>
<tr>
<td>CI</td>
<td>confidence interval</td>
</tr>
<tr>
<td>CVD</td>
<td>cardiovascular disease</td>
</tr>
<tr>
<td>ECG</td>
<td>electrocardiography</td>
</tr>
<tr>
<td>FI</td>
<td>Frailty Index</td>
</tr>
<tr>
<td>HRQL</td>
<td>health-related quality of life</td>
</tr>
<tr>
<td>kB</td>
<td>kilo base</td>
</tr>
<tr>
<td>LTL</td>
<td>leukocyte telomere length</td>
</tr>
<tr>
<td>MCID</td>
<td>minimal clinically important difference</td>
</tr>
<tr>
<td>PA</td>
<td>physical activity</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>RCT</td>
<td>randomized controlled trial</td>
</tr>
<tr>
<td>RAND-36</td>
<td>RAND-36–Item Health Survey</td>
</tr>
<tr>
<td>RT-qPCR</td>
<td>real-time quantitative polymerase chain reaction</td>
</tr>
<tr>
<td>SD</td>
<td>standard deviation</td>
</tr>
<tr>
<td>SE</td>
<td>standard error</td>
</tr>
<tr>
<td>SF-36</td>
<td>MOS-36–Item Short Form Health Survey</td>
</tr>
<tr>
<td>TL</td>
<td>telomere length</td>
</tr>
<tr>
<td>TRF</td>
<td>terminal restriction fragment</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
List of original articles

This thesis is based on the following publications, which are referred to throughout the text by their Roman numerals:


Contents

Abstract
Summary in Finnish
Acknowledgements 9
Abbreviations 13
List of original articles 15
Contents 17
1 Introduction 19
2 Review of the literature 21
   2.1 Defining physical activity .............................................................. 21
   2.2 Measuring physical activity in epidemiologic studies ...................... 22
      2.2.1 Objective physical activity measures ........................................ 22
      2.2.2 Subjective physical activity measures ....................................... 23
   2.3 Physical activity recommendations ............................................... 25
      2.1 Socioeconomic factors and physical activity .................................. 26
      2.2 Health-related quality of life (HRQL) ........................................... 27
         2.2.1 Definition of HRQL ............................................................... 27
         2.2.2 Measuring HRQL ................................................................. 28
         2.2.3 Physical activity and HRQL .................................................. 32
   2.3 Old age frailty ............................................................................ 35
      2.3.1 Aetiology and pathogenesis of frailty ....................................... 36
      2.3.2 Definitions of frailty ............................................................... 38
      2.3.3 Frailty assessment tools ......................................................... 39
      2.3.4 Physical activity and prevention of frailty .................................. 42
      2.3.5 Physical activity and treatment of frailty ................................... 43
      2.3.6 Comorbidity, disability and sarcopenia in the context of frailty ... 44
   2.4 Biomarkers of ageing ................................................................. 45
   2.5 Telomeres .................................................................................. 46
      2.5.1 Telomere length and ageing ...................................................... 47
      2.5.2 Measuring telomere length ...................................................... 48
      2.5.3 Physical activity and telomere length ....................................... 50
   2.6 Physical activity in the context of mortality and CVD risk factors .... 52
      2.6.1 Mortality ............................................................................. 52
      2.6.2 Physical activity and CVD risk factors .................................... 53
3 Purpose of the research 57
4 Subjects and methods 59
   4.1 Subjects ................................................................................................................. 59
   4.2 Flowchart of the study ................................................................................................. 61
   4.3 The voluntary examinations in 1964–1973 ................................................................. 62
   4.4 Baseline examination in 1974 and physical activity assessment .............. 62
       4.4.1 Physical activity assessment in 1974 ................................................................ 63
   4.5 Physical activity surveys in 1986 and 2000 ................................................................. 63
   4.6 Questionnaire surveys in 2000 for HRQL and frailty status ......................... 64
       4.6.1 HRQL assessment ............................................................................................. 64
       4.6.2 Frailty assessment ............................................................................................. 65
   4.7 Telomere length measurement ................................................................................. 65
   4.8 Mortality follow-up ................................................................................................. 66
   4.9 Statistical methods ................................................................................................. 66
5 Results 69
   5.1 HRQL and physical activity ...................................................................................... 71
   5.2 Frailty and physical activity ...................................................................................... 72
   5.3 Telomere length and physical activity ...................................................................... 74
   5.4 Mortality and physical activity .................................................................................. 75
6 Discussion 79
   6.1 HRQL and physical activity ..................................................................................... 80
   6.2 Frailty and physical activity ..................................................................................... 81
   6.3 Telomere length and physical activity ...................................................................... 83
   6.4 Mortality and physical activity .................................................................................. 85
   6.5 Strengths and limitations ......................................................................................... 87
       6.5.1 Strengths ............................................................................................................ 87
       6.5.2 Limitations of the physical activity assessment and categorization ................ 87
       6.5.3 Limitations of the cohort and settings ............................................................. 88
   6.6 Conclusions ............................................................................................................ 90
References 93
Original articles 121
1 Introduction

All parts of the body, if used in moderation and exercised in labours, to which each is accustomed, become thereby healthy, well developed and age slowly; but if they are unused and left idle, they become liable to diseases, defective in growth and age quickly.

Hippocrates
Greek physician, fifth century B.C

Physical activity (PA) has been strongly and consistently shown to promote longevity (Samitz et al. 2011) and to prevent several diseases (Haskell et al. 2007, Lee et al. 2012, Warburton et al. 2006). However, leading a life with little or no PA has been made increasingly feasible by rapid industrialization and the development of modern technology (Lagerros & Lagiou 2007). Hence, despite all the knowledge of the beneficial effects of PA, physical inactivity has emerged as a major health problem worldwide, particularly among adolescents and older adults (Hallal et al. 2012). In 2009, the World Health Organization (WHO) stated physical inactivity as one of the most important risk factors for global mortality (World Health Organization 2011).

During the past 30 years, population has been ageing especially in high-income countries, due to both higher life expectancy and lower birth rate. Between 1995 and 2010, the proportion of adults aged 65 years or older increased from 12% to 15% in OECD countries (OECD 2013). In 2013, the proportion of people in Finland aged 65 years or older was as high as 19% (Statistics Finland 2014).

In the face of the broad societal challenges following the current ageing trends, scientists and policy makers have focused on identifying factors that may have a positive impact on the ageing process, age-related diseases as well as the quality of life of older adults. Regular PA has been proposed to be one such promising area. (King & King 2010). PA has already been taken as a feasible and relatively inexpensive way to strive against non-communicable diseases (Berryman 2012). Although it is evident that PA has a positive influence on ageing and health, many questions remain unanswered (Lagerros & Lagiou 2007). The maturation of cohort studies has made it possible to investigate the long-term associations of PA with various aspects of ageing.
2 Review of the literature

2.1 Defining physical activity

In the health sciences, the widely used definition of PA is by Caspersen from 1985. He defined PA as any bodily movement produced by skeletal muscles that results in energy expenditure (Caspersen et al. 1985). PA can be categorized in a variety of ways, for example according to type, intensity, duration, or purpose. Different types include aerobic PA, muscle strengthening and balance improvement (Troiano 2009). A commonly used categorization is based on the context in which PA is performed, often called different domains of PA, for example occupational, leisure-time, recreational, transportation etc. Many aspects of PA can be differentiated either to discretionary or nondiscretionary (Troiano et al. 2012).

Leisure-time PA is defined as PA which is not required as essential activities of daily living and is performed by discretion. These activities may include sports participation, exercise training as well as recreational activities, such as going for a walk, dancing and gardening (Physical Activity Guidelines Advisory Committee 2008). Exercise is a subcategory of PA having the same elements as PA, but it is also “planned, structured, and repetitive”, with improvement or maintenance of physical fitness as its objective. The intensity of PA can be defined and categorized in many different ways, e.g. according to the rate of energy expenditure (e.g. kilocalories per minute, or multiples of resting energy expenditure [MET]), or in some cases simply by the speed of the activity (walking 3 miles per hour). It can also be classified into categories such as low, moderate, vigorous or hard and very hard.

Physical fitness is a distinct entity, which has been defined, among several other definitions, as a set of either health- or skill-related attributes, e.g. cardiorespiratory endurance, skeletal muscle endurance, flexibility, balance and speed of movement (Caspersen et al. 1985).

The term sedentary behaviour has conventionally been used to describe behaviour which does not include moderate-to-vigorous intensity exercise. While recent epidemiologic evidence shows that the consequences of habitual sedentary behaviour (i.e., too much sitting) are distinct from the consequences of lack of moderate-to-vigorous exercise (Dunstan et al. 2010), non-exercise behaviour has been recommended to be differentiated into two categories: sedentary behaviour.
2.2 Measuring physical activity in epidemiologic studies

At the early stages of PA epidemiology, the focus of interest was especially on occupational PA (Morris et al. 1953, Paffenbarger & Hale 1975). Since the PA level at work has been decreasing for the last 50 years (Church et al. 2011), leisure-time PA measurements have been assumed to be the best representation of the PA in the population (Caspersen et al. 1985, Kriska & Caspersen 1997). Recently, the interest in work-related PA has been increasing again (Church et al. 2011). Today, the energy expenditure at work varies substantially within populations, and detailed information about occupational and commuting PA is valuable in many respects. Furthermore, as interest in the area of sedentary behaviour has increased, there has been an effort to assess low intensity activities and sedentary behaviour more comprehensively (Helmerhorst et al. 2012).

Measuring PA levels of large populations is challenging, as PA itself is a multidimensional and complex construct. The optimal PA assessment tool should measure all the domains of PA, including intensity, duration and frequency. In addition, it should be convenient to participants and researchers, and have good validity and reliability.

The choice of a PA assessment tool, or a combination of them, depends on several characteristics of the study, e.g. whether the PA information is used to control confounding by PA or to explore the effects of different PA domains (Sternfeld & Goldman-Rosas 2012). All measurements have their limitations, but can be useful in particular instances. PA measurements can be divided into objective (device-based) and subjective (report-based) methods.

2.2.1 Objective physical activity measures

Doubly labelled water technique, which is often called a gold standard of PA measurements, and the respiratory chamber are suggested to be the most exact objective quantitative assessments of total energy expenditure (Lagerros & Lagiou 2007, LaPorte et al. 1985). During the last decade, measures calculating movement, such as pedometers, accelerometers and GPS(global position system)-based equipment, as well as measures estimating some parameters of physical fitness, like maximal oxygen uptake measures and heart rate monitoring, have
become feasible even for large epidemiological studies. For example, accelerometers have been used in cross-sectional studies of 6,329 Americans (Troiano et al. 2008), 4,696 Portuguese (Baptista et al. 2012), 3,267 Norwegians (Hansen et al. 2012) and 1,589 Finns (Husu et al. 2014). Numerous studies have used subjective measures for the entire study population and objective measures in a smaller sample to validate the subjective measurement into their study population (Aadahl et al. 2007, Emaus et al. 2010, Friedenreich et al. 2006).

It has been shown that self-reported PA level correlates to physical fitness (Sandvik et al. 1993), which is easier to measure objectively than total energy expenditure. Thus, physical fitness has often been considered as an objective measurement of PA, or at least as a more accurate measure than self-reported PA (Williams 2001). It should be kept in mind that physical fitness may reflect inherited physiological and metabolic characteristics of individuals (Lee et al. 2012) and is not as such a direct, objective measure of PA.

The objective PA measures have their own limitations. The more exact methods, such as doubly labelled water and respiratory chambers, are not so prone to information bias, but are time-consuming and costly laboratory tests (Lagerros & Lagiou 2007). Devices that are used in large populations, such as accelerometers and pedometers, may measure well a certain type of exercise, but may not count other types (Troiano et al. 2012). For example, a waist-worn accelerometer may quantify well horizontal locomotion like walking and running, but not bicycling, swimming or strengthening exercises. Moreover, most of the currently available devices provide little or no information about the type, purpose or relative intensity of PA for a given individual (Troiano et al. 2012), all of which are important aspects of habitual PA.

2.2.2 Subjective physical activity measures

Subjective ways to assess PA include self-reports such as records (diary type), logs (a list of specific types of PA to choose from), recalls (typically collected by interviewers using memory-enhancing questions) and questionnaires.

Physical activity questionnaires (PAQs)

In epidemiological studies PA is typically assessed by PAQs (Paffenbarger et al. 1993, Samitz et al. 2011, Slattery et al. 1989, Talbot et al. 2007, Yu et al. 2003), which vary from a single question (Steptoe et al. 2002) to very detailed
questionnaires (Craig et al. 2003). In 2011, Samitz and colleagues reviewed 80 cohort studies about PA and mortality, and all the eligible studies used PAQs as PA assessment tools (Samitz et al. 2011). Approximately half of the studies used one to four questions or a brief global PA assessment, while the other half used detailed PAQs.

The advantage of PAQs is their practicability: they are convenient for participants, they can be designed to suit the particular population, they do not alter the behaviour of the individual being surveyed, as records and logs may do, and the costs are reasonable (Lagerros & Lagiou 2007). Furthermore, PAQs can be designed to directly answer the research question.

**Limitations of PAQs**

The main weakness of PAQs is that they are susceptible to information bias and misclassification (Lagerros & Lagiou 2007, Tooze et al. 2013) as well as to social desirability bias, which refers to a defensive tendency to answer in a manner that is consistent with social norms or beliefs (Adams et al. 2005, Klesges et al. 2004, Sallis & Saelens 2000). The overestimation of PA is a common type of bias (Banda et al. 2010, Matton et al. 2007, Troiano 2009). This bias may be differential, as self-reported PA is suggested to be more accurate at a high level of PA (Aadahl et al. 2007, Emaus et al. 2010, Friedenreich et al. 2006, Jacobs et al. 1993).

PAQs vary a lot in their complexity, and recalling and reporting of PA may be a demanding task (Baranowski et al. 1998). Especially children and very old adults may have insufficient cognitive capacity in completing PAQs (Durante & Ainsworth 1996).

The over-reporting and thus over-estimation of PA has been considered to cause attenuation of the true effect of PA in studies relying on PAQs (Hallal et al. 2012, Matton et al. 2007, Troiano 2009). This is supported by a recent study, which examined the validity of a comprehensive PAQ with 433 participants and compared the PAQ answers to PA level assessed by double-labelled water. (Tooze et al. 2013).

**Reliability and validity of PAQs**

The reliability of PAQ is usually evaluated by the test-retest procedure to measure the consistency of PAQs. High reliability indicates that the instrument generates
the same outcome when repeated under the same conditions. High validity of PAQ refers to the instrument’s ability to measure exactly what it is supposed to measure. (Kriska & Caspersen 1997). The validity of PAQ should be estimated against objective PA assessment, such as accelerometer or doubly labelled water technique.

In a recent systematic review of the reliability and objective-criterion related validity of 130 PAQs, only very few questionnaires had both good reliability and validity. The PAQs mostly had acceptable reliability, and PAQs developed after 1999 had slightly better reliability than older ones, while validity was poor in almost all questionnaires. Across all PAQs, sedentary behaviour was poorly correlated to objective measures of PA. (Helmerhorst et al. 2012).

Validation of PAQ in one population does not evidently mean it can be generalized to other populations, especially to other age groups (Corder et al. 2009). In a review of 13 PAQs designed for the elderly, the methodology of the validation procedure was often found to be inadequate, and none of the PAQs were found to have good validity and reliability using high-quality methods (Forsen et al. 2010).

To improve the quality of PAQs and their validation studies, new recommendations have been presented (Hagstromer et al. 2012) and measurement error models developed (Tooze et al. 2013).

Currently, the most accurate method to assess PA in large-scale studies has been suggested to be the simultaneous use of a subjective and objective method, e.g. PAQ and accelerometer (Haskell 2012, Kirk & Rhodes 2011, Troiano et al. 2012).

2.3 Physical activity recommendations

The global PA recommendation for adults is 150 minutes of moderate intensity aerobic PA per week or 75 minutes of vigorous PA per week, or an equivalent combination of moderate and vigorous PA, and muscle strength or balance training at least two times a week. For additional health benefits, adults are recommended to increase their aerobic PA to 300 minutes of moderate intensity or 150 minutes of vigorous intensity PA (Haskell et al. 2007). These recommendations have been poorly followed: worldwide, 31.1% of people have been estimated to be inactive (Hallal et al. 2012). In the United States, 48% of adults meet the recommended minimum if PA is assessed by self-report, which is known to overestimate the true level of PA, but only 5% meet the recommended
minimum level if PA is assessed by accelerometers (Troiano et al. 2008). In Norway, 20%, and in Portugal, 7–9% of adults meet the recommended PA level if PA is assessed by accelerometer (Baptista et al. 2012, Hansen et al. 2012). In Finland, in a survey relying on self-reported PA only 14% of the population aged 20–55 met the PA recommendations for both aerobic PA and muscle strengthening while 38% met the recommendations partly (National Institute for Health and Welfare 2014). In a recent survey using accelerometers as PA assessment 24% of 1,589 Finnish participants aged 18-85 met the PA recommendations for aerobic exercise (Husu et al. 2014). In most countries, men meet PA recommendations more often than women and younger adults more often than older adults (Hallal et al. 2012).

PA issues are becoming increasingly important in Public Health policy and amendments (Berryman 2012). In 2011, the United Nations stated PA as one of the cornerstones in the prevention of non-communicable diseases, e.g. type 2 diabetes, cardiovascular diseases (CVD) and breast and colon cancers (General Assembly of the United Nations 2011).

2.1 Socioeconomic factors and physical activity

Globally, there are more physically inactive people in high-income countries than in low-income countries (Hosseinpoor et al. 2012). In an analysis of 66 countries, 45% of the population in high-income countries was found to be inactive while 24% of the population in low-income country was inactive (Hallal et al. 2012). Over the last decades, occupational and commuting PA, which are typically more common in people with low income, have been decreasing worldwide (Brownson et al. 2005, Knuth & Hallal 2009). There is pronounced concern about the increase of physical inactivity in developing countries and the consequent increase of the burden of noncommunicable diseases. To prevent and delay this tendency, WHO has published the Global Recommendations on Physical Activity for Health, which is primarily targeted to policy-makers on national level (World Health Organisation 2010).

Within countries, socioeconomic factors are likely to have multifaceted and country-specific effects on PA levels and trends (Bauman et al. 2011, Kim et al. 2004). Lack of time and motivation are more common barriers to engaging in PAs among people with high income, while illness, disability, lack of money and transportation are more common barriers among people with low income (Chinn et al. 1999). Furthermore, the domains of PA, amount of stress (Krueger & Chang
and risk factors for noncommunicable diseases (Hosseinpoor et al. 2012) vary in their prevalence and modifiability in different socioeconomic groups.

Studies that rely solely on leisure-time PA estimates have suggested that PA is more common in high socioeconomic class than in low socioeconomic class (Trost et al. 2002). However, in several studies, the total PA has been found to be higher among people with lower socioeconomic status than among people with higher socioeconomic status (Kirk & Rhodes 2011, Lee et al. 2007). This may indicate that the people in high socioeconomic status occupations may not be accumulating enough PA during leisure time to compensate for their sedentary occupation and are at risk of insufficient activity (Kirk & Rhodes 2011). Overall, the relationships between occupation and PA are very complex, and there are transitions at individual and population levels (Hillsdon 2011).

2.2 Health-related quality of life (HRQL)

2.2.1 Definition of HRQL

Health has traditionally been measured narrowly, often by life expectancy, mortality and morbidity statistics (Drewnowski & Evans 2001). In 1946, the WHO defined health broadly as: “A state of complete physical, mental and social well-being and not merely the absence of disease or infirmity” (World Health Organization 1946). Since then, there has been steadily more interest in patient-oriented outcomes: how a patient feels, physically and psychologically, how he/she gets along with other people and copes with everyday life (Sullivan 2003), and since 1970 the use of the term quality of life has increased nearly exponentially (Testa & Simonson 1996).

Quality of life is defined by the WHO as: “Individuals’ perceptions of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns”. In social and medical literature, quality of life has a wide range of descriptions varying from life satisfaction to a board multidimensional construct (Rejeski & Mihalko 2001).

The concept of HRQL has evolved since the 1980s to cover those aspects of quality of life that are associated with physical or mental health (McHorney 1999). There are several definitions of HRQL in literature, and they often have the following core dimensions: physical, social and psychological dimensions;
global assessment of perceived health may also be included (Aalto Anna-Mari et al. 1999).

In gerontology, HRQL has been increasingly in the focus. In 1995, the Gerontological Society of America formulated the following motto: “Adding life to years, not just more years to life”.

2.2.2 Measuring HRQL

In all HRQL assessments, the subjective perception, i.e., patients’ own first-hand view, is the prerequisite for translating the degree of patients’ health into actual experienced quality of life (Testa & Simonson 1996). The HRQL measurements can be classified into two main genres: generic and disease-specific measurements (Hays & Morales 2001).

Generic measures

Generic measures are intended to be appropriate for groups with various health statuses and co-morbidities, and they are usually comprehensive, covering several domains of health and quality of life. These measures enable comparisons between different populations, interventions and treatments.

Generic measures can be further divided into profile measures and preference-specific measures. Profile measures are designed to yield scores on multiple domains of HRQL (Hays & Morales 2001). The most popular generic measures include the Sickness Impact Profile (Bergner et al. 1981), the Nottingham Health Profile (Hunt et al. 1985) and the RAND-36–Item Health Survey (RAND-36), which is described below in more detail.

Preference-specific scales include weighted measures of health state, which offer a single aggregate score to be used particularly in economic estimations. These scales are often also referred to as utility assessments. Quality-Adjusted Life Years (QALYs) is a widely used concept developed for cost-utility studies. A QALY means a year of full life quality (Weinstein & Stason 1977, Weinstein et al. 2009), and its estimation presumes both life expectancy and HRQL scores.

Disease- and population-specific measures

Disease- and population-specific measures are designed to be used in specific patient populations, such as patients with cancer, arthritis or depression. There are
dozens of measures of different clinical conditions and diseases, and for many
diseases, numerous different scales exist (Garratt et al. 2002).

Dimension-specific measures are used when some dimension (like physical
functioning or cognition) is of particular interest. These measures are not always
generated to be HRQL measurements, and they are often used as supplement to
disease-specific or generic measures. An example of this is Beck’s depression
inventory (Beck et al. 1961).

LEIPAD is a population-specific measure developed for older adults (De Leo
et al. 1998).

**RAND-36**

RAND-36 is a generic profile HRQL measure, which was originally developed
under the name MOS-36 in the United States in the late 1980s as part of the
Medical Outcome Study, a large longitudinal investigation of patients with a
range of chronic conditions. The 36 items (questions) of this HRQL instrument
are selected from a larger pool of items used in the study. The “standard” form of
the instrument was first available in 1990 (Ware & Gandek 1998). This 36-item
battery has been distributed with different names by different organizations: the
MOS-36–Item Short Form Health Survey (SF-36) by the Medical Outcomes
Trust, RAND-36 by the RAND organization, the Health Status Questionnaire by
the Health Outcomes Institute, and the RAND-36 Health Status Inventory by the
Psychological Corporation.

The nearly identical RAND-36 and SF-36 instruments have similar wording,
but minor deviations in scoring for pain and general health scales. It has been
shown that these differences between the English versions of RAND-36 and SF-
36 have minimal effect on the scale scores (Hays et al. 1993). However, the
outcomes from some translated versions of RAND-36 and SF-36 are not fully
comparable with each other (Aaronson et al. 1998). The use of SF-36 requires a
permission fee, while RAND-36 is freely available from the RAND organization.
The RAND-36/SF-36 questionnaire has been translated for use in more than 60
different countries (International Quality of Life Assessment Project 2014) and it
is the most widely used generic HRQL measure in the world (Garratt et al. 2002,
Hays & Morales 2001).
The content of the RAND-36 questionnaire

The RAND-36 includes 36 items (questions) that cover eight dimensions (scales) of HRQL:

1. physical functioning (ten questions)
2. role limitations caused by physical health problems (four questions)
3. bodily pain (two questions)
4. general health perceptions (five questions)
5. vitality (four questions)
6. role limitations caused by emotional problems (three questions)
7. social functioning (two questions)
8. mental health/emotional well-being (five questions)
9. reported change in health (rating of health now, compared to one year ago)

The first four scales (1–4) make up the physical health summary score and the following four scales (5–8) the mental health summary score. The last question (9) about change in health is not included in the scoring of the eight scales or summary scores.

The response choices to these questions vary from two (dichotomous) to six. The scoring of the scales includes two steps. First, the answers to each item are linearly transformed to a 0 to 100 scale, 100 indicating the highest (most favourable) and 0 the lowest level of function and well-being. Next, the scores to scales are calculated by summing the scores of items in the scale, and dividing it by the number of items in the particular scale. If there are missing answers, the score in the scale can be calculated from the answers to the other items of that scale, if there are at least three items in that scale and at least 50% of the items in that scale have been properly completed (Aalto Anna-Mari et al. 1999).

The two summary scales, i.e., the physical and mental health summary scores, are further calculated from the eight scales of RAND-36. They are standardized so that the mean score for the validated population is 50. The summary scales are not validated in the Finnish population (Aalto Anna-Mari et al. 1999).

Reliability and validity of RAND-36/SF-36

The reliability of the eight scales and two summary scores of RAND-36/SF-36 was first investigated in the population of the MOS study in terms of internal
consistency and test-retest reliability. Internal consistency refers to the ability to get the same outcome when measuring the same concept by different items. In case of multi-item measures it also refers to the fact that the sum of the items gives more information than a single item. Test-retest reliability refers to the ability to get the same outcome on different occasions for the same population, when there is no evidence of change.

The reliability coefficient has been found to range from 0.65 to 0.94 across scales, indicating sufficient reliability for group comparisons (Bland & Altman 1997, Jenkinson et al. 1994, McHorney et al. 1994).

The content validity, the ability to cover all the relevant aspects of HRQL, has been tested against other generic health surveys (Ware & Gandek 1998). RAND-36/SF-36 has been shown to be sensitive to changes in 36-month follow-up (Hemingway et al. 1997).

The predictive validity of RAND-36/SF-36 has been studied in several prospective studies and both physical and mental summary scores have been shown to predict mortality (Fan et al. 2002, Kroenke et al. 2008, Tsai et al. 2007). RAND-36 has also been shown to predict hospitalization among veterans (Fan et al. 2002).

Floor effects (tendency to get a substantial amount of the lowest score of the scale) have been shown for the two role limitations scales. Ceiling effects (opposite to floor effect) were observed for both role disability and the social functioning scales (McHorney et al. 1994). The proportion of missing values in questionnaires has been 1.1–5.9% (McHorney et al. 1994).

In 1995, the RAND-36/SF-36 was translated into Finnish, and it has been studied for its reliability and validity. The Finnish version of RAND-36 has been found to be a reliable and valid measure of HRQL in the Finnish general population (Aalto Anna-Mari et al. 1999). Numerous other studies have been conducted on the reliability and validity of the RAND-36/SF-36 instrument after translating the questionnaire into different languages. Overall, the instrument has been found to be a sufficiently valid and reliable instrument for measuring HRQL.

Minimal clinically important difference

The concept of minimal clinically important difference (MCID) was developed to refer to the smallest difference in outcomes which is clinically significant,
worthwhile and important. It has been suggested that the MCID for the RAND-36 would be in the range of 3 to 5 points (Samsa et al. 1999).

In the context of HRQL measurements, the identification of a single score for MCID is very complex and changes below the suggested MCID should not be ignored, but considered thoroughly (Hays & Woolley 2000).

Limitations of RAND-36

Although it only takes approximately 7 to 10 minutes to complete the RAND-36 (Hays & Morales 2001), the participation rates in postal questionnaire surveys have been around 70% (Djarv et al. 2013, Jenkinson et al. 1994) This may cause conservative bias, as non-respondents are likely to be from lower socioeconomic class and likely to have lower RAND-36 scores (Hemingway et al. 1997, Parker et al. 2006). As self-report, the RAND-36 is susceptible to reporting bias, such as social desirability bias (Klesges et al. 2004, Sallis & Saelens 2000).

To reach feasibility, some important contents of HRQL are not included in RAND-36, such as sleep adequacy, cognitive functioning, family functioning, self-esteem or health distress. RAND-36 is a multi-item instrument with its own benefits, but the responsiveness for minor changes in a single item may be attenuated by scores from other items (Patterson 2000).

2.2.3 Physical activity and HRQL

There is growing evidence from cross-sectional and longitudinal epidemiologic studies, as well as from intervention studies, that PA and exercise are associated with better HRQL at all ages. The improvements are most often found in physical function, vitality and other domains of physical health, but in some studies, improvements are also found in domains of mental health. The measurements of HRQL and PA vary across presented studies, but most studies have used RAND-36/SF-36 and PAQs. Few studies have used objective measures of PA (accelerometers); in those studies, the associations between PA and HRQL have been stronger (Anokye et al. 2012, Wanderley et al. 2011).

Intervention studies on exercise and HRQL are often conducted in certain subgroups, e.g. cancer survivors or CVD patients, but there are some RCTs in general populations. Improvements in all eight dimensions of HRQL in a dose-responsive manner and independently of weight change were found after a 6-month exercise intervention among 430 sedentary postmenopausal women.
In a trial of 151 men and women at moderate-to-high CVD risk, those in the intervention group (attending supervised exercise sessions and dietary counselling during 3 months, followed by group meetings for up to 3 years) had an improved sense of physical functioning, less bodily pain, and better vitality and social functioning than the control group (Eriksson et al. 2010). That study was conducted in primary health care without any additional resources, and was found to be highly cost-effective as the intervention group had less visits to the family physician.

In a trial of 76 community-dwelling older women pronounced improvements in physical functioning, vitality and general health were observed after a 25-week adapted PA programme, but no significant improvement in mental health scores were seen (Kovacs et al. 2013). In an exercise intervention study conducted in elderly people right after hospital admission, the intervention group had better scores in vitality and bodily pain than the control group. The intervention group performed progressive resistance training consisting of five exercises: flexion and extension of the hip, knee and leg and sit-to-stand exercise. (Brovold et al. 2012)

In contrast to the findings of positive associations between PA and HRQL, no improvement was found in HRQL in an RCT of 202 middle-aged men and women after a 12-month exercise intervention consisting of 360min/week of moderate-to-vigorous aerobic exercise. In subgroup analyses only overweight male exercisers got higher 12-month HRQL scores in role physical, vitality and social functioning than the controls (Imayama et al. 2011) Furthermore, there was no statistically significant improvement in QL, assessed by WHOQOL-Bref questionnaire after a 16-week resistance training trial in 16 postmenopausal women compared to control group (Bonganha et al. 2012), which may be due to the short duration of the intervention and the small number of participants. In an RCT of 149 older home-dwelling women, an exercise programme intended to reduce the risk of falls and fractures had only a slight effect on HRQL; only the general health score improved during the 12-month survey, but this gain was lost at 24 months (Karinkanta et al. 2012).

In a comprehensive meta-analysis of 53 physical exercise RCTs, a small but meaningful improvement in quality of life was found for 3–6 months in healthy populations and also in patients exercising as part of their rehabilitation. However, there was a small deterioration among patients involved in exercise interventions as part of their treatment or management of chronic conditions or health events. (Gillison et al. 2009).
In a recent review and meta-analysis, older adults with functional limitations and co-morbidity were found to improve their mobility and physical function by exercise interventions, although no consistent improvements in QOL were found (Vries et al. 2012). In 2001, Rejeski and colleagues reviewed 18 studies, mostly intervention studies, on PA and HRQL among older adults. The majority of participants had some chronic disease, most commonly CVD. PA and HRQL were measured variably in the studies, but it was concluded that PA could have positive effects on both physical functioning and mental health status among older adults (Rejeski & Mihalko 2001). In their earlier review, the findings about the associations of PA and QL were more equivocal (Rejeski et al. 1996).

There are several cross-sectional epidemiologic studies that have found positive associations between PA and HRQL, usually stronger associations in physical functioning and other physical domains of HRQL (Anokye et al. 2012, Franco et al. 2012, Herman et al. 2012, Kruger et al. 2007, McNaughton et al. 2012, Moilanen et al. 2012, Pucci et al. 2012, Wanderley et al. 2011). Nevertheless, some studies have also found significant positive associations between PA and mental health (Baernholdt et al. 2012, Sanchez-Villegas et al. 2012, Vallance et al. 2012, Vathesatogkit et al. 2012). In a recent population-based cross-sectional analysis of 4,190 participants in Sweden, poor HRQL was strongly associated with inactivity. The most important covariate to HRQL was reported diseases, and the second strongest covariate was physical inactivity (Djarv et al. 2013). In general adult populations, the association between PA and HRQL has been shown to be stronger among people older than 65 years (Bertheussen et al. 2011) and independent of weight change (Herman et al. 2012, Kruger et al. 2007). Furthermore, an analysis of 35,425 subjects showed an inverse association between physical fitness and complaints of QL in a dose-response fashion (Mitchell & Barlow 2011).

In a review of healthy subjects under 65 years of age, there was a consistent association between higher levels of PA and higher scores in various HRQL dimensions in all seven cross-sectional studies. The association was most pronounced with physical functioning and vitality (Bize et al. 2007). This review included a large cross-sectional analysis of 175,850 people, which showed that meeting the recommended levels of PA is associated with better HRQL and self-perceived health (Brown et al. 2003).

Recent longitudinal studies on PA and HRQL have concluded that there exists a positive association between PA and both the physical and mental domains of HRQL. In a study of 1,097 older adults, PA level during 6 years of follow-up was
significantly associated with better scores in all eight domains of HRQL (Balboa-Castillo et al. 2011). In a study of 3,621 participants, persistently active subjects reported highest scores in almost all domains of HRQL, and subjects increasing their PA reported higher scores in physical function, vitality and general health than subjects remaining inactive or with decreasing PA (Oostrom et al. 2012). Similarly, in a study of 63,152 women, a clear increase in PA during 10 years of follow-up was associated with improvements in all eight domains of HRQL (Wolin et al. 2007), and in a study of 4,206 participants, maintenance or increase in PA during 8-year follow-up were associated with better scores in HRQL, especially in mental domains (Sanchez-Villegas et al. 2012). In an analysis of 18,344 women, high level of PA as well as walking were associated with better HRQL in a curvilinear manner. The associations with mental health scores attenuated during the 6-year follow-up (Heesch et al. 2012). In a follow-up study of 13,535 women, midlife PA was strongly associated with both physical function and mental health index, which were included in their definition of successful survival (Sun et al. 2010a, Sun et al. 2010b).

### 2.3 Old age frailty

Frailty is a relatively new concept, which has its roots in the geriatric literature of the 1980s and 1990s (Abellan van Kan et al. 2008). Today, the term is increasingly used in medicine. It has been described as a biological syndrome of decreased reserves and impaired stress tolerance, resulting from cumulative declines across multiple physiologic systems leading to vulnerability to adverse outcomes (Fried et al. 2001, Lang et al. 2009b). Frailty may occur without a clear culprit or as a result of a various diseases and medical conditions (Morley et al. 2013), and the terms primary frailty and secondary frailty have been used to describe these states, respectively (Abellan van Kan et al. 2008, Fried et al. 2004, Strandberg & Pitkala 2007). Frailty is not an inevitable consequence of ageing (Strandberg et al. 2011, Abellan van Kan et al. 2008) and about 73% of people older than 85 years do not meet the criteria for frailty (Collard et al. 2012).

The concept of frailty has been constructed to identify older adults at an increased risk of adverse health outcomes such as disabilities, falls, loss of independence, hospitalization, institutionalization and increased mortality (Fried et al. 2001). Frailty has been found to be an independent risk factor for adverse postoperative outcomes (Dasgupta et al. 2009, Makary et al. 2010). It has been considered that preoperative frailty screening would be valuable in risk

According to a recent systematic review incorporating 21 cohorts, the mean prevalence of frailty among community-dwelling persons aged 65 and older varies between 10% and 14%, depending on the frailty definition used. Prefrailty, referring to the state between robustness and frailty, has been observed in as many as 41.6%. The prevalence of frailty in women is nearly twofold compared to men, and the prevalence increases substantially in age groups older than 80 years in both sexes. (Collard et al. 2012). The higher prevalence of frailty is also related to lower socioeconomic class, higher rates of comorbid diseases and disability (Fried et al. 2001). Among elderly patients undergoing elective surgery, the prevalence of frailty has been reported to be between 42% and 51% (Afilalo et al. 2009, Makary et al. 2010, Sundermann et al. 2011), which reflects the vulnerability of frail people (Partridge et al. 2012).

The different frailty models classify older adults as frail with slightly different sociodemographic profiles, but generally frail adults are older, more likely to be female, unmarried, from a minority ethnic group, and more likely to live alone, have less education and a lower net worth than those not classified as frail (Cigolle et al. 2009). Furthermore, elderly people living in deprived neighbourhoods (Lang et al. 2009a), having little contact with relatives or neighbours, not participating in community or religious activities or in helping others (Woo et al. 2005) are more likely to be frail as classified by deficit accumulation models of frailty (Lang et al. 2009a).

2.3.1 Aetiology and pathogenesis of frailty

The aetiopathology of the physical phenotype of frailty is multifactorial, encompassing deficits in several inter-related physiologic systems. The complex mechanisms of ageing are involved in this cumulative decline and the subsequent depletion of homeostatic and physiologic reserves (Clegg et al. 2013). Clinical frailty occurs when a threshold is reached (Fried et al. 2009). The possible aetiological factors of frailty include genetic and epigenetic factors (Kirkwood 2005, Kohl 3rd et al. 2012) as well as environmental and lifestyle factors (Bergman et al. 2007). Also traumas (Fried et al. 2001) and atherosclerosis (Newman et al. 2001) may be involved.
Declines in cognitive and neuronal systems, sarcopenia, anorexia, inflammation and certain hormones, such as corticosteroids, have been proposed to be some plausible pathogenetic factors in frailty (Abellan van Kan et al. 2008, Clegg et al. 2013, Fried et al. 2009, Strandberg et al. 2011).

An overview of the aetiopathology, possible interventions and consequences of the physical phenotype of frailty is presented in Figure 1.

Fig. 1. Schematic representation of frailty and PA.
2.3.2 Definitions of frailty

The concept of frailty has been present in geriatric literature at least since 1972 when Stamford and colleagues used the term synonymously with institutionalization (Stamford 1972). The more operational definitions of frailty were developed in the 1980s and 1990s (Abellan van Kan et al. 2008). After that, there have been several proposals of frailty with different emphases, varying from specific subsets to broad concepts including cognitive, functional and social circumstances. Recently, interest in disease- or even procedure-specific frailty concepts has awakened among physicians of varying specialities (Iqbal et al. 2013).

The definitions of frailty can be divided into two main models: physical frailty phenotype and deficit accumulation models of frailty. The unresolved distinctions between these main models have prevented reaching a successful consensus of frailty issues, such as an international definition. In 2013, the consensus group of delegates from six major international, European and American societies came into the conclusion that physical frailty is a specific syndrome within the broader context of frailty (Morley et al. 2013).

Physical phenotype of frailty

In 2001 Fried and colleagues presented a qualitative, operational definition of frailty, often called the physical phenotype of frailty or Fried Criteria (Fried et al. 2001), which is today the most widely used definition of frailty in clinical and epidemiologic research (Bouillon et al. 2013, Collard et al. 2012). This definition is based on data from the Cardiovascular Health Study (Fried et al. 2001) and Women’s Health and Ageing Studies (Bandeen-Roche et al. 2006). The definition conceptualizes frailty as a biological syndrome of decreased resiliency, in which both known and unknown root causes result in clinically measurable manifestations (Strandberg et al. 2011). Nonphysical domains, such as social, cognitive and psychological domains, have been excluded from the diagnostic criteria as distinct, albeit highly important, entities (Abellan van Kan et al. 2010). The physical frailty diagnosis is based on the constellation of five possible components: unintentional weight loss, subjective exhaustion, weakness, slowness and reduced PA (Fried et al. 2001).
In the expanded models of physical frailty, nonphysical components such as cognitive and psychological items have been added to the definition of frailty (Abellan van Kan et al. 2010).

The physical phenotype has been shown to be associated with specific alterations in physiological variables, including inflammatory biomarkers, altered glucose metabolism and markers of clotting processes (Walston et al. 2002), supporting the view of multisystemic involvement.

**Deficit accumulation models of frailty**

In 1994, Rockwood and colleagues presented a definition of frailty, which is based on accumulation of identified deficits in several areas of health, later called the frailty index (FI) (Rockwood et al. 1994). These deficits encompass cognitive status, mood, motivation, communication, mobility, balance, bowel and bladder function, nutrition and social resources, activities of daily living as well as a number of comorbidities. The Canadian Study of Health and Ageing was used in the elaborating surveys (Rockwood et al. 1994). The same cohort was used by Mitnitski and colleagues when they developed the assessment tool further (Mitnitski et al. 2001). Several other FIs have also been developed (Bouillon et al. 2013).

The essential distinction between the physical phenotype and deficit accumulation models of frailty is that phenotypic frailty may be distinct from disability and comorbidities as a biological syndrome with a pathophysiology of its own, while FIs include a wide range of deficits regardless of their nature (Strandberg et al. 2011).

**2.3.3 Frailty assessment tools**

Two recent reviews of existing frailty instruments included 20 and 27 different instruments (Bouillon et al. 2013, Vries et al. 2011), which illustrates the variety of frailty assessment tools. Currently, numerous tools meet most of the validity criteria suggested for frailty instruments (Morley et al. 2013): content validity (i.e., includes multiple determinants, is dynamic and useful across contexts), construct validity (i.e., more common with age and in women, related to disability, comorbidity and self-rated health) and criterion validity (i.e., predicts mortality and other adverse outcomes) as well as sufficient reliability (Rockwood 2005). Choosing a frailty assessment tool depends on several aspects: is the
purpose clinical or research; is the assessment done for risk stratification or case-finding, and is the assessment meant to be carried out by a research assistant, geriatrician or public health physician.

**Fried’s Scale**

Fried’s Scale and its modifications have been used most widely in clinical and epidemiologic studies. The frailty status is assessed by the following clinical characteristics (Fried et al. 2001):

1. unintended weight loss (loss of $\geq 10$ pounds or $\geq 5\%$ of body weight in the past year),
2. weakness (grip strength in the lowest 20%),
3. exhaustion (self-report of fatigue or feeling unusually tired or weak in the past month)
4. slowness (walking 4m $\geq 7$ s if height $\leq 159$ cm or 6 s if height $\geq 159$ cm)
5. low PA (frequency and duration of PAs, lowest 20%).

If three or more criteria are met, the person is classified as frail; if one to two criteria are met, the person is classified as pre-frail; and if none of the criteria are met, the person is classified as robust (Fried et al. 2001).

In various studies, modifications have often been made to the criteria depending on the data available.

(Chang et al. 2011), cognition (Kiely et al. 2009), disability (Kiely et al. 2009), chronic medical conditions (Kiely et al. 2009) and HRQL (Masel et al. 2009). Internal construct validity (Bandeen-Roche et al. 2006) and convergent validity (Rockwood et al. 2007) have also been tested. Reliability has not been reported for Fried’s Scale, in line with most frailty assessments (Bouillon et al. 2013). All things considered, Fried’s Scale can be considered as a well-validated frailty assessment tool for the physical phenotype of frailty.

Other physical frailty assessment tools

Numerous other physical frailty instruments exist, varying from expanded methods of Fried’s Scale to very simple assessments. A few examples are given in the following.

The assessment of cognitive function has been added to the Fried Scale, and that has been shown to improve its predictive value (Avila-Funes et al. 2009). The “FRAIL” screening tool includes a question on severe comorbidity (more than 5 diseases) in addition to more physical items (Morley et al. 2012). Some quite simple, user-friendly tests have also been shown to be sufficient. The SOF index includes three components: weight loss, inability to rise from chair 5 times without using arms and reduced energy level. SOF index has been shown to predict adverse outcomes, such as falls, disability and deaths (Ensrud et al. 2008). According to Gill and colleagues, persons are considered physically frail if they require more than 10 seconds to perform a rapid-gait test (i.e., to walk along a 3-m course and back as quickly as possible) or if they are unable to stand up from a seated position in a hardback chair with their arms folded (Gill et al. 1995). Gait speed at the usual pace has been suggested to be alone a predictor of vulnerability and poor outcomes in older adults (Abellan van Kan et al. 2009b). In a recent pooled analysis of 9 cohort studies, median life expectancy was predicted at a speed of 0.8m/s on a 4-m course, and a cut-off speed of 0.6m/s was recommended for identifying persons with significantly shorter than average survival (Studenski et al. 2011).

Self-reported frailty assessment tools without detailed clinical measures, such as gait speed, have been considered to be well suited in certain situations (Martin & Brighton 2008). The Groningen Frailty Indicator is a screening instrument consisting of 15 self-reported items that focus on physical, social, psychological and cognitive domains of frailty (Steverink et al. 2001). The Tilburg Frailty Indicator includes two subscales. The first subscale is composed of determinants
Frailty Index

A common way to calculate FI is to count the prevalent deficits and divide the sum by screened deficits (Mitnitski et al. 2002, Rockwood et al. 2005). The provided estimate varies between 0 (absence of deficits) to 1 (presence of all screened deficits). The cut point of 0.25 has been used for frailty and 0.08 for pre-frailty, and a seven-level categorization of frailty has also been used (Song et al. 2010a). The number of measured deficits in validated FIs usually varies between 30 and 70 (Song et al. 2010a). The deficits may be co-morbidities, symptoms, signs or laboratory abnormalities, and they reflect impairments in numerous domains of health including cognition, mood, communication, mobility, balance, nutrition and social resources. In a traditional FI, 70 deficits have been screened, which makes it time-consuming and burdensome. Hence, Rockwood and colleagues have developed a Clinical Frailty Scale to serve as a quick multidimensional frailty assessment tool (Rockwood et al. 2005). They have also presented a standard procedure for selecting appropriate deficits and creating FI from an ageing database (Searle et al. 2008). The validity of different FIs has been studied from various aspects. The predictive (Mitnitski et al. 2002, Rockwood et al. 2005, Rockwood et al. 2007, Song et al. 2007), concurrent (Cigolle et al. 2009) and construct (Mitnitski et al. 2002) validity as well as inter-rater validity (Rockwood et al. 2005) has been tested and found to be sufficient or good.

2.3.4 Physical activity and prevention of frailty

As frailty is a relatively new concept, the research on frailty prevention is limited to date. Higher prevalence of frailty is associated with e.g. subclinical
atherosclerosis, CVDs (Afilalo et al. 2009), obesity (Blaum et al. 2005), hyperglycaemia (Blaum et al. 2009), Alzheimer’s disease (Buchman et al. 2007) and co-morbidity (Fried et al. 2004). It could be suggested that these diseases share the same root causes as frailty and are involved in the pathogenesis (Strandberg et al. 2012). As healthy lifestyle habits, including PA, prevent several diseases (Sofi et al. 2011, Warburton et al. 2006), a healthy lifestyle at all ages most probably prevents frailty as well.

The primary prevention of frailty is obviously a life course approach (Bergman et al. 2007, Strandberg et al. 2011). This view is supported by a few observational studies. Midlife obesity and cardiovascular risk factors (Strandberg et al. 2012) as well as PA in older adulthood (Peterson et al. 2009) have been shown to be associated with old-age frailty. Furthermore, PA in midlife has been shown to have a strong protective effect on mobility disability in later life (Fries et al. 1994, Paterson & Warburton 2010) and falls (Sherrington et al. 2008). Resistance training is the most efficacious way to prevent and reverse sarcopenia (Abellan van Kan et al. 2009a), which is closely linked to frailty. A recent meta-analysis showed that PA prevents and slows down the disablement process in aged or diseased populations (Tak et al. 2013).

In addition to lifestyle factors, the adequate and prompt treatment of diseases and traumas of elderly people are the cornerstones of frailty prevention and treatment. However, there is no evidence for pharmacological therapy in the primary prevention of frailty (Strandberg et al. 2011).

### 2.3.5 Physical activity and treatment of frailty

Currently, the best evidence from observational studies and RCTs in the treatment of frailty is for physical exercise (Strandberg et al. 2011). Systematic reviews including RCTs have found exercise to have positive impacts on muscle function, cardiorespiratory function, flexibility and functional ability in both institutionalized and community-dwelling frail older adults (Chou et al. 2012, Daniels et al. 2008, Theou et al. 2011, Weening-Dijksterhuis et al. 2011). In recent RCTs on frail community-dwelling older adults, an exercise programme prevented progression of frailty and was cost-effective among 610 participants (Yamada et al. 2012), while a nutrition and exercise programme improved frailty stage in 3-month follow-up, but not in 6- or 12-month follow-up among 117 participants (Chan et al. 2013). A multifactorial intervention including exercise was found to improve physical status significantly compared to the control group.
in a 12-month survey, but not in a 3-month survey in a trial of 240 participants (Cameron et al. 2013), which supports the view that the time frame needed for positive changes in frailty stage is year(s) rather than a few months.

For the management of frailty state, 45 minutes of exercise 3 times a week has been recommended (Morley et al. 2013, Theou et al. 2011). The exercise should include progressive resistance training, balance training and functional training (Weening-Dijksterhuis et al. 2011).

There is some evidence for other possible treatments for frailty in addition to physical exercise: caloric and protein supplementation (Malafarina et al. 2013, Milne et al. 2009, Tieland et al. 2012), which is most effective when combined with exercise training and vitamin D supplementation to frail persons, who are vitamin D deficient (Muir & Montero-Odasso 2011). Furthermore, the management of reversible diseases has to be included in the treatment of frail older adults (Morley et al. 2013). In later and end-stage frailty, the main goal is good quality of life and palliation (Strandberg et al. 2011).

2.3.6 Comorbidity, disability and sarcopenia in the context of frailty

The associations between comorbidity, disability, sarcopenia and frailty are complex, as illustrated in Figure 1. They overlap with each other, and sometimes these entities have been confused.

Comorbidity, defined as having two or more disorders or diseases, is a common phenomenon in frail people. Chronic diseases are involved in the aetiology and pathogenesis of frailty (Newman et al. 2001, Strandberg et al. 2011) and frailty itself is an independent risk factor for the development and progression of certain diseases, such as CVDs (Afilalo et al. 2009). Thus, they form a vicious circle where one factor exacerbates the other. On the other hand, proper treatment of diseases may prevent or revert frailty, and treating frailty may relieve diseases. (Iqbal et al. 2013). Nevertheless, multimorbidity is more widespread than frailty, affecting 3 out of 4 persons older than 65 years (Tinetti et al. 2012), and it is evident that the majority of elderly people with comorbidity are not frail; in the landmark study of Fried and colleagues, more than 90% of the participants with comorbidity were not frail (Fried et al. 2001).

Disability, which is usually defined as difficulty in at least one activity of daily living, often occurs synergistically with frailty, similarly to comorbidity (Fried et al. 2004). As frailty progresses to its later stages, disability is always present (Strandberg et al. 2011). One major aim of frailty case finding is to find
frail pre-disabled individuals as it allows interventions that could prevent disability (Morley et al. 2013). In this respect, it would be highly important to distinguish disability from frailty syndrome (Abellan van Kan et al. 2008). Comorbidity and disability are included in several frailty scales, such as FIs (Bouillon et al. 2013).

The substantial overlapping of disability, co-morbidity and frailty has been reported in several studies: 68–84% of frail participants have been reported to have comorbidity (Fried et al. 2001, Theou et al. 2012, Wong et al. 2010). Two studies reported that 27–29% of frail participants had disability (Fried et al. 2001, Wong et al. 2010), while one study reported that as many as 84% of frail participants had disability and only a very small proportion, 4–8% of frail participants, did not have either diseases or disability (Theou et al. 2012).

Sarcopenia is a syndrome characterized by progressive and generalized loss of muscle mass combined with low muscle function (Cruz-Jentoft et al. 2010). Although frailty and sarcopenia overlap, frailty is a more multifaceted syndrome than sarcopenia (Bauer & Sieber 2008, Cruz-Jentoft et al. 2010, Morley et al. 2013). Sarcopenic obesity, which refers to loss of muscle mass while fat mass is preserved or increased (Cruz-Jentoft et al. 2010), is also a common component of frailty, leading to “fat-frail” people. This type of frailty is more challenging to recognize, as we intuitively identify frail people as thin and shrunken (Strandberg et al. 2011). As both obesity and inactivity are emerging health problems, it is most probable that sarcopenic obesity as part of frailty syndrome is becoming even more common.

2.4 Biomarkers of ageing

Biomarkers of ageing can be defined as parameters which would predict functional capacity, age-related diseases and mortality better than chronological age (Boonekamp et al. 2013, Sprott 2010). The American Federation for Aging Research has proposed the following criteria for a biomarker of ageing: it must predict the rate of ageing; monitor a process central to biology of ageing, instead of a specific disease; be convenient to be tested repeatedly and accurately without harming the person; and work both in humans and in laboratory animals (Johnson 2006). Other criteria also exist in the literature (Butler et al. 2004, Sprott 2010). Several parameters such as inflammatory cytokines, corticosteroids and telomere length (TL) have been examined from this point of view (Johnson 2006, Mishra et al. 2012), but currently none of them fully meets the strict criteria and high
requirements for biomarker of ageing (Falandry et al. 2013). Nevertheless, TL has been adopted as a remarkable hallmark of ageing (Lopez-Otin et al. 2013) and it is widely considered to be a marker of biological ageing in literature (Fyhrquist et al. 2013).

2.5 Telomeres

Telomeres are nucleoprotein structures located at the ends of chromosomes. They consist of tandem repeats of TTAGGG DNA sequences and a number of highly specialized proteins. In human telomeres, the TTAGGG sequence is repeated more than several thousand times (Moyzis et al. 1988). Elizabeth Blackburn, Jack Szostak and Carol Greider, who were all awarded the Nobel Prize for telomere research in 1991, found that telomeres protect chromosomes from erosion during cell division and that a specific enzyme, telomerase, is involved in the preparation and elongation of telomeres. (Blackburn 1991, Greider & Blackburn 1985, Szostak & Blackburn 1982).

In cell cultures, telomeres are shown to shorten by 50–100 base pairs (BP) with each cell division (Harley et al. 1990, Proctor & Kirkwood 2002). When a critically short TL is reached, the chromosomes are left “uncapped”, leading to senescence (i.e., nondividing state of the cell) and apoptosis (d’Adda di Fagagna et al. 2003), but in less severe cases damage repair and “re-capping” of chromosomes may occur (Blackburn 2000). Telomere erosion explains the limited proliferative capacity of some types of in vitro cultured cells, the so-called replicative senescence or Hayflick limit (Hayflick & Moorhead 1961, Lopez-Otin et al. 2013, Olovnikov 1996).

In humans, the telomerase enzyme is active especially in germline, embryonic and stem cells, and in smaller amounts also in several types of proliferating somatic cells, for example in the immune system, skin and hair follicles (Blackburn 2000). However, most mammalian somatic cells do not express telomerase (Lopez-Otin et al. 2013).

In humans, leukocyte telomere length (LTL) at birth varies from about 5 kilo bases (kB) to 15 kB and it is for the most part heritable (Broer et al. 2013, Hunt et al. 2008, Jeanclos et al. 2000, Slagboom et al. 1994). LTL is associated with race (Hunt et al. 2008) and gender (Diez Roux et al. 2009), with males losing telomeres faster than females, and the loss decelerating throughout the human lifespan (Unryn et al. 2005). In Caucasian adults (aged 35–55), mean LTL typically varies between 7 kB and 8.5 kB (De Meyer et al. 2007) and the annual

2.5.1 Telomere length and ageing

The data from epidemiologic studies supporting the inverse association between age and TL are quite robust. In a recent review of 124 cross-sectional and 5 longitudinal studies, there was a clear inverse correlation between the mean age and the mean LTL (Muezzinler et al. 2013). In the cross-sectional studies with more than 1,000 participants, telomere loss rate varied mostly between 20–30BP/year. All the 5 longitudinal studies found an inverse association between TL and age. Three of the studies provided estimates of annual telomere loss rate, which ranged from 32 to 46 BP/year, being somewhat higher than the loss rates found in cross-sectional studies. Interestingly, all longitudinal studies also reported telomere elongation in a small subset of their participants. Whether this is a true biological elongation over time or consequence of measurement errors is not yet clear. (Muezzinler et al. 2013).

TL and ageing-related diseases

The associations between TL and diseases have been studied intensively during the last decades. Short telomeres are causally involved in certain rare hereditary diseases by mutations in genes coding telomerase components, such as dyskeratosis congenita (Vulliamy et al. 2008), a disease characterized by premature ageing, some types of aplastic anaemia (Yamaguchi et al. 2005) and pulmonary fibrosis (Alder et al. 2008, Armanios et al. 2007). On the contrary, tumour cells typically express high levels of telomerase, which prevents telomere shortening (Harley & Kim 1996).

Several studies have reported positive associations between shortened LTL and CVD (Benetos et al. 2004, Samani et al. 2001, Samani & van der Harst 2008), CVD events (Brouilette et al. 2007), stroke (Ding et al. 2012, Jiang et al. 2013) and subclinical atherosclerosis (Panayiotou et al. 2010). Among 419 participants of the Cardiovascular Health Study, LTL was associated with factors related to CVD, and among participants under 73 years, there was three-fold increased risk of myocardial infarction or stroke with every 1 kB shorter LTL (Fitzpatrick et al. 2007). In a prospective study of 800 participants with 10-year follow-up, shorter LTL was strongly associated with CVD risk (Willeit et al. 2007).
2010). In contrast, some studies have found no associations between LTL and stroke (Fyhrquist et al. 2011, Zee et al. 2010). Although conclusive evidence is lacking, telomere attrition is suggested to have an important role in CVD (De Meyer et al. 2011, Fyhrquist et al. 2013).

Researchers have reported positive associations between shorter LTL and cognitive decline (Valdes et al. 2010, Yaffe et al. 2011), but also negative associations with physical and cognitive decline have been reported (Harris et al. 2006). Furthermore, LTL has been found to be related to diabetes (Sampson et al. 2006, Zee et al. 2010) and cancer (Willeit et al. 2010).

**TL and mortality**

An inverse association between LTL and mortality has been found in several studies, but the correlation between LTL and survival declines among older people (Boonekamp et al. 2013, Cawthon et al. 2003, Harris et al. 2006, Martin-Ruiz et al. 2005). In an analysis of 2,721 participants, LTL was not found to be associated with mortality; instead, LTL was significantly correlated to self-reported years of healthy life (Njajou et al. 2009).

**2.5.2 Measuring telomere length**

In epidemiological studies, TL is usually measured from DNA of peripheral blood leukocytes, as blood samples are convenient and feasible for large populations. TL varies in different tissues of an individual, but it has been suggested that LTL could serve as a surrogate parameter for the relative TL of all tissues (Friedrich et al. 2000). A single critically short telomere has been shown to lead the cell to senescence (Abdallah et al. 2009, Bendix et al. 2010, Hemann et al. 2001) and researchers have increasingly estimated also the proportion of very short telomeres.

There are several methods to measure TL, and two most commonly used methods in epidemiology are described here. It is notable that different methods result in different estimates, and the number of subtelomeric regions in the estimates varies between methods. Furthermore, the measurement errors of the same method may differ between laboratories (Aviv et al. 2011). Comparisons between methods and studies thus have to be made with caution.
Terminal restriction fragment (TRF) analysis

A quantitative method to measure mean TL by TRF analysis was reported in the late 1980s (Moyzis et al. 1988) and it is still a widely used technique to measure TL. TRF analysis has been considered as a “gold standard” for TL measuring techniques, and it has been used to validate and calibrate most new techniques (Aubert et al. 2012). TRF analysis provides an estimate of the average TL of the cell population.

TRF analysis has some important advantages compared to other TL measurement techniques: TRF analysis can measure TL distribution, e.g. the proportion of very short telomeres, the coefficient variation between study groups is less than 2% (Kimura et al. 2010), and it provides estimates in kB or nucleotides.

The major drawback of the TRF technique is that it requires a considerable amount of DNA: 0.5–5.0 ug DNA or genomic DNA from at least $10^5$ cells (Cawthon 2002), and it is costly and labour-intensive (Aubert et al. 2012). Furthermore, TFR measures both the canonical region of the telomere (strictly TTAGGG repeats) and non-canonical region of the telomere, which reaches up to the nearest restriction site, which is specific for the particular enzymes used. Thus, the mean TFR lengths may vary as much as 5% according to the enzymes used (Cawthon 2002), affecting comparability between studies using different enzymes. Also, the restriction site polymorphism and the length of the subtelomeric region cause TFR length variations between individuals. In studies comparing the TL sizes measured by TFR and QPCR or Q-FISH, subtelomeric regions have been estimated to be approximately 2–4kB (Cawthon 2002, Counter et al. 1992, Steinert et al. 2004), and variation between participants to be 2 kB (Cawthon 2002).

Real-time quantitative polymerase chain reaction (RT-QPCR) analysis and monochrome multiplex QPCR (MM-QPCR)

In 2002, Cawthon presented the real-time quantitative polymerase chain reaction (RT-QPCR) method to measure TL (Cawthon 2002). Seven years later, a further improved version, monochrome multiplex quantitative PCR (MM-QPCR), was developed (Cawthon 2009). The method is based on the expression levels of the telomere sequence (target) and single copy gene (reference gene). The ratio of the
telomere copy number (T) to single copy gene (S) is referred to as the T/S ratio and it is proportional to the average TL.

The main advantages of the RT-QPCR and MM-QPCR techniques against other TL measurements are low cost, high throughput and the small amount of DNA required (50ng per sample). The disadvantages are that only the average TL is measured and the coefficient of variation between groups is more than 2%, and the lack of good reference standards makes absolute TL measurement difficult (Kimura et al. 2010).

2.5.3 Physical activity and telomere length

TL changes slowly, over years (Ludlow et al. 2013), which poses challenges to intervention studies. Thus, most of the current information about the relationship between PA and TL comes from cross-sectional studies. Although a growing body of evidence suggests that higher levels of PA are associated with longer LTL, the reported associations between PA and TL are not entirely consistent, especially at the higher levels of PA.

In 2008, Cherkas and colleagues showed that among 2,401 twin volunteers aged 18–81, there was a positive association of LTL with increasing self-reported leisure-time PA level (Cherkas et al. 2008). The mean LTL of the most active subjects was 200 BP longer than the mean LTL of inactive subjects, which can be suggested to be 5–8 years’ difference in biological age, when assuming an LTL loss rate of 25–40 BP per year (Muezzinler et al. 2013). In their sub-analysis of 15 monozygotic and 52 dizygotic twin pairs who were currently discordant for PA, there was an 88 BP difference in LTL, when there was an at least 2-point difference in PA level when using a 4-point scale (Cherkas et al. 2008). Similar results were found in a study of 44 healthy postmenopausal women (mean age 58), in whom habitual PA was related to longer LTL. (Kim et al. 2012). Furthermore, in a large cross-sectional analysis of 7,813 women in the Nurses’ Health Study, moderately or vigorously active women had longer mean LTL than less active women. The longest LTL were found among women engaging in moderate or vigorous activity 2.5 hours per week, but the researchers did not observe additional benefits from longer duration or higher intensity of PA in terms of LTL (Du et al. 2012).

In some studies conducted at high levels of PA, the differences found in favour of exercise are remarkable. Seventeen endurance-exercise trained adults with a mean age of 63 had 900 BP longer LTL than their sedentary controls.
(LaRocca et al. 2010), and in a recent study, 57 marathon runners were found to have 10% longer telomeres than their healthy controls, which was suggested as a 16-year difference in chronological age (Denham et al. 2013). Interestingly, in a trial of 9 moderately trained males, longer TL in circulating T-lymphocyte cells was detected after a single exhaustive exercise bout (Simpson et al. 2010).

In a study of 944 CVD patients aged 53–82, low exercise capacity was strongly associated to shorter LTL, while self-reported physical inactivity was associated to shorter LTL only in unadjusted models. (Krauss et al. 2011). The fact that self-reported PA was not associated to LTL after multivariable adjustment, whereas objectively assessed exercise capacity was, could be due to the bias caused by self-report.

An inverse u-shaped correlation between PA levels and LTL has been found among 69 participants aged 50–70. Moderate levels of PA were related to higher LTL compared to the lowest and highest PA level (Ludlow et al. 2008). In that study, no association was found between telomerase activity and PA level.

Some researchers have not observed a direct association between PA and LTL; instead, associations have been found either in clusters of lifestyle habits or in subgroups. In a study of 612 prostate cancer cases and 1,049 controls with a mean age of 63, there was no significant association between PA and TL; instead, a positive association was found between a cluster of healthy living habits (exercise, tobacco, body mass index [BMI] and diet) and longer TL (Mirabello et al. 2009). Changes in lifestyle habits including PA led to LTL lengthening in a 5-year follow-up of 35 prostate cancer patients (Ornish et al. 2013). In a study of 63 postmenopausal women aged 54–82, PA level was not directly associated to LTL, but among participants reporting high perceived psychological stress exercise prevented LTL shortening (Puterman et al. 2010).

There are some studies that have not found an association between PA and LTL. In an intervention study of 234 postmenopausal women aged 50–70 years, exercise or exercise and diet did not lead to a change in LTL in 12-month follow-up (Mason et al. 2013). Nevertheless, at the baseline, LTL was positively associated to maximal oxygen uptake. The absence of change in LTL could be explained by too short a follow-up time, as the change in LTL may become slowly detectable. In a cross-sectional analysis of 2,006 participants aged 65 years and over (Woo et al. 2008) and 2,284 women with a mean age of 59 (Cassidy et al. 2010) PA was not found to be related to LTL. In a study of 80 participants, PA was not significantly associated to LTL; instead, PA was observed to correlate to cellular biomarkers of telomere dysfunction (Song et al. 2010b). In an
investigation of 17 marathon runners, the runners did not have longer LTL than their 15 age-matched controls. The runners practiced on average 51 km/week and the average duration time was 14 years, which can be considered as vigorous exercise for several years (Mathur et al. 2013).

Some researchers have investigated the relationship between TL in skeletal muscle and PA. Collins and colleagues reported that athletes with exercise-associated fatigue had shorter muscle TL than the controls (Collins et al. 2003). In a study of 10 healthy endurance athletes and 10 controls, endurance training history was associated with longer muscle TL (Osthus et al. 2012), while in some other studies no association has been observed between muscle TL and resistance training (Kadi & Ponsot 2010) or endurance training (Laye et al. 2012, Rae et al. 2010).

2.6 Physical activity in the context of mortality and CVD risk factors

In 2009, WHO stated physical inactivity as the fourth leading risk factor for global mortality after high blood pressure, high glucose and tobacco use, even before overweight and obesity (World Health Organization 2011).

In several large-scale cohort studies, PA has been associated to lower risk of developing CVDs (Boone-Heinonen et al. 2009, Gordon-Larsen et al. 2009, Lee et al. 2012, Williams 2001), which are the most substantial cause of mortality worldwide. In 2011, 30.4% of the deaths worldwide and even 48.3% of the deaths in Europe were caused by CVD, meaning over 16 million deaths per year.

2.6.1 Mortality

PA has been shown to be inversely associated with all-cause mortality in several cohort studies (Andersen et al. 2000, Chakravarty et al. 2012, Leon et al. 1987, Manini et al. 2006, Moore et al. 2012, Paffenbarger et al. 1986.). A recent meta-analysis of 80 prospective epidemiologic studies with altogether 1,338,143 participants yielded an estimation of 29% reduction in all-cause mortality between the lowest and highest levels of PA in adjusted models (Samitz et al. 2011). The number of covariates in those studies varied from 2 to 23, with a median of 7 variables, most often age, smoking, BMI and blood pressure. In minimally adjusted models, the risk reduction was as high as 46% (Samitz et al. 2011).
A dose-response-like association between PA and mortality has been shown in several studies (Paffenbarger et al. 1986, Wen et al. 2011) and has been further supported by recent dose-response meta-analyses (Janssen et al. 2013, Samitz et al. 2011), thus enhancing the message “some is good, more is better”. This relation appears to be curvilinear, so that engaging in PA 150 min per week gives about 20% risk reduction in all-cause mortality compared with PA engagement of less than 30 min/week. Additional amounts of PA give more risk reduction, but at slightly smaller magnitudes, i.e., 7 hours per week gives about 40% risk reduction in mortality (Janssen et al. 2013). It is notable that even amounts under the recommend levels of PA give a mortality risk reduction compared to totally inactive people (Rockhill et al. 2001, Wen et al. 2011).

An association between higher mortality and lower socioeconomic status, as measured by income, education or occupation, has been shown in several epidemiologic studies over the decades (Doornbos & Kromhout 1990, Guralnik et al. 1993, Laaksonen et al. 2008, Lin et al. 2003, Liu et al. 1982, Salonen 1982, Sorlie & Rogot 1990). This difference in mortality has been partly explained by differences in health-related behaviour and material factors, but it is not yet fully understood. The impact of health-related behaviours on differences in mortality has been suggested to be around 50%, and the most important behaviours have typically been smoking and inactivity (Laaksonen et al. 2008, Woodward et al. 2003).

### 2.6.2 Physical activity and CVD risk factors

PA has been shown to reduce the most important CVD risk factors, high blood pressure, obesity, hypercholesterolaemia and high blood glucose, in several cohort studies and in RCTs. Hence, the mortality reduction by PA is partly modified through CVD risk factors (Johansson & Sundquist 1999).

#### High blood pressure

Both aerobic (Cornelissen & Fagard 2005, Whelton et al. 2002) and resistance exercise (Cornelissen et al. 2011) have been shown to reduce resting blood pressure, the biggest risk factor for mortality worldwide (World Health Organization 2011). According to a large meta-analysis of 72 RCTs regular PA reduced resting systolic blood pressure by 3 mmHg and resting diastolic blood pressure by 2.4 mmHg on average, while the blood pressure reduction was even
6.9/4.9 mmHg in hypertensive participants (Cornelissen & Fagard 2005). Hypertension treatment guidelines recommend regular PA as the first-line non-pharmacological treatment for hypertension (Mancia et al. 2013). The Guidelines of the European Society of Cardiology suggest that hypertensive patients should be advised to engage in moderate intensity aerobic exercise for 30 minutes 5–7 times per week. Dynamic resistance exercise 2–3 times per week could also be advised. PA has also been adopted as one of the most important prevention strategies for hypertension (Mancia et al. 2013).

**Obesity**

There is clear evidence that habitual PA is associated with prevention of excess weight gain in both women and men (Bottai et al. 2014, Thune et al. 1998), and this is evident even in old age (Di Pietro et al. 2004). There is a dose-response relation between PA and weight loss, yet dietary intervention is usually needed as energy balance is always dependent on both energy intake and expenditure (Ross & Janssen 2001).

**Hypercholestrolaemia**

The relationship between regular PA and dyslipidaemia has been widely studied, and the elevation of HDL cholesterol has been the main finding. According to meta-analyses (Kodama et al. 2007, Leon & Sanchez 2001), regular PA consuming 900–1,200 kcal per week or more results in a modest but significant HDL cholesterol elevation. In a review of 25 RCTs, this significant elevation was described as 2.53 mg/ml or 0.065 mmol/l, and at the authors’ rough estimate, this HDL elevation would reduce CVD risk by 5% for men and 8% for women (Kodama et al. 2007). PA has been found to be beneficial in reducing triglycerides by the same amounts of exercise as needed for the elevation of HDL (Slentz et al. 2007). LDL cholesterol has also been found to decrease in some PA interventions (Kodama et al. 2007, Leon & Sanchez 2001).

**High blood glucose and diabetes**

PA has a clear, important role in preventing and treating high levels of blood glucose. The evidence from high-quality RCTs shows that both resistance and aerobic PA can lead to improved insulin action lasting from 2 hours to 72 hours...
and can assist in managing blood glucose levels (Colberg et al. 2010). In an RCT conducted in 577 people with high risk of diabetes, a sole exercise intervention reduced the risk for developing diabetes by 46%, while a diet intervention reduced the risk by only 31% and a diet-plus-exercise intervention by 42% over a 6-year period (Pan et al. 1997). PA and diet are usually combined in RCTs, lowering the risk for diabetes. In a Finnish RCT with 522 subjects, a diet and exercise intervention reduced the risk for diabetes by 58% over 3.2 years of follow-up (Tuomilehto et al. 2001) and similar results were found in an RCT conducted in 3,234 subjects in the US (Knowler et al. 2002).

Large prospective cohort studies show that an increase in regular PA decreases the risk of high blood glucose and type 2 diabetes (Jeon et al. 2007) and is also advantageous in the control and cure of type 2 diabetes (Sigal et al. 2004).
3 Purpose of the research

The purpose of these analyses was to study the long-term relationships of self-reported PA in midlife among 782 males of the Helsinki Businessmen Study. The specific aims were:

1. to investigate the associations of PA to HRQL in old age (I)
2. to investigate the associations of PA to frailty status in old age (II)
3. to investigate the association of PA to LTL and the proportion of short telomeres in old age (III)
4. to investigate the associations of PA in midlife to mortality, independently of cardiovascular risk factors, including BMI, age, cholesterol, glucose, systolic blood pressure and smoking at baseline (IV)
4 Subjects and methods

The study population is part of the Helsinki Businessmen Study, a cohort gathered in the 1960s and 1970s at the Institute of Occupational Health in Helsinki, Finland. The follow-up study procedures including baseline examinations, questionnaires and laboratory tests have been approved by the Ethics Committee of the Department of Medicine, University of Helsinki.

4.1 Subjects

In 1964–1973, 3,490 men born in 1919–1934 had voluntarily participated in structured health check-ups at the Institute of Occupational Health in Helsinki, Finland. Initially, the health check-ups were not performed for scientific purposes, but the socioeconomic homogeneity of the participants was noted as an appropriate basis for research and further proceedings were made to find healthy middle-aged men for a primary prevention study. For scientific use, data are available on 3,313 men, 95% of the 3,490 men who took part in the health check-ups in 1964–1973. Most of the participants were business executives or had academic education, a few were sole traders or higher office workers. Approximately 60% of the participants were from Helsinki or the Uusimaa region, the rest were from different parts of Finland. At that time, occupational health check-ups were not customary in Finland.

The 3,313 men were evaluated with questionnaires and clinical and laboratory examinations in 1974, whereupon 1,815 men were found to be actively working and healthy without diabetes, clinical CVD or regular medications. By 1974, 68 men out of the 3,313 had died; 867 men refused to take part and 563 men were excluded from the study.

The exclusion criteria in 1974 were: medical treatment for hypertension, CVD, diabetes or hyperlipidaemia; hypertension (systolic blood pressure $\geq 200$ mmHg and/or diastolic blood pressure $\geq 115$ mmHg) or secondary hypertension; CVDs such as history of myocardial infarction, angina pectoris, ECG changes, history of clinical cardiomyopathy, valvular disease or heart failure; cerebrovascular diseases such as history of TIA or stroke (vertebral haemorrhage or ischemic stroke), unilateral symptoms or findings of hemiparesis; renal failure (S-creatinine $\geq 150$\,umol/l) or renal diseases; metabolic diseases such as diabetes (fasting glucose $\geq 10$mmol/l); malignant diseases; psychiatric diseases and alcoholism.
Table 1. Baseline characteristics of 1,815 participants of the Helsinki Businessmen Study and 563 men who were excluded, 867 men who refused to participate and 68 men who had died by 1974.

<table>
<thead>
<tr>
<th>Variable in baseline</th>
<th>Participants (n=1815)</th>
<th>Excluded (n=563)</th>
<th>Refused (n=867)</th>
<th>Dead by 1974 (n=68)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in 1974, years</td>
<td>48 (4)</td>
<td>49 (4)</td>
<td>48 (4)</td>
<td>50 (4)</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>26 (3)</td>
<td>26 (3)</td>
<td>26 (3)</td>
<td>26 (3)</td>
</tr>
<tr>
<td>Blood pressure, mmHg</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systolic</td>
<td>133 (13)</td>
<td>144 (20)</td>
<td>136 (17)</td>
<td>147 (22)</td>
</tr>
<tr>
<td>Diastolic</td>
<td>85 (8)</td>
<td>92 (13)</td>
<td>86 (10)</td>
<td>93 (15)</td>
</tr>
<tr>
<td>Cholesterol, mmol/l</td>
<td>7.1 (1.2)</td>
<td>7.4 (1.3)</td>
<td>7.2 (1.3)</td>
<td>7.6 (1.4)</td>
</tr>
<tr>
<td>Smokers, %</td>
<td>37</td>
<td>46</td>
<td>47</td>
<td>52</td>
</tr>
<tr>
<td>Keys Risk score</td>
<td>18 (14)</td>
<td>29 (28)</td>
<td>21 (22)</td>
<td>43 (50)</td>
</tr>
</tbody>
</table>

Data are means, standard deviation (SD) in parenthesis, unless otherwise stated.

A structured questionnaire about details of their PA was available for 782 of the 1,815 men. In 1974, this detailed information was mostly (90.7%, n = 709) gathered from men who had at least one of the following risk factors: high cholesterol level, high serum triglyceride level, high systolic or diastolic blood pressure, smoking more than ten cigarettes per day, overweight, or impaired glucose tolerance. These men were assessed to be at higher CVD risk and were studied more extensively in order to find healthy participants with CVD risk factors for the initial primary prevention study. However, their age distribution or long-term mortality was not statistically different from the men without detailed PA assessment.

Although part of the men participated in the primary prevention trial during the 1970s, preliminary analyses showed that participation in the 1970s trial did not affect the long-term results. Therefore, all the 782 men with detailed PA information form the population of these analyses. In the studies (I-IV), the number of subjects varies depending on the methods used. The mean age of the participants was 47.8 (SD 4.1, range 40–58) at the baseline in 1974.
4.2 Flowchart of the study

Fig. 2. Flowchart.

1964 - 1973
3490 Businessmen born in 1919-1934, participated in structured health check-ups at the Institute of Occupational Health in Helsinki.
Data available, n = 3314

1974
Questionnaires, clinical examination and laboratory tests
Healthy, without diabetes, clinical CVD or regular medications n=1815

1974
Mean age 48
Detailed questionnaires about their PA
n = 782

1985 - 1986
Mailed questionnaires, n = 631

2000
Mean age 73
Mailed questionnaire including RAND-36
Study I, HRQL assessment, n = 552
Study II, Frailty assessment, n = 514

2002 - 2003
Mean age 70
Randomly selected survivors for laboratory tests, LTL analysis
Study III, n = 204

2007
Mortality follow-up, cumulative death between 1974 and 2007, n = 295
Study IV, n = 782
4.3  The voluntary examinations in 1964–1973

The voluntary health examinations at the Institute of Occupational Health in Helsinki were performed on 3,490 men. The primary aim of these examinations was to provide occupational health care, not actually to gather data for scientific use. The participants were evaluated by a physician, and height, weight and blood pressure were measured and smoking status was assessed. Laboratory tests were taken and serum cholesterol, triglycerides and blood glucose one hour after a glucose load (1g/kg of body weight of glucose orally) were determined. After the year 1973, the examinations were repeated every 2 to 5 years until 1979–1980.

4.4  Baseline examination in 1974 and physical activity assessment

A mailed questionnaire and an invitation to ECG and laboratory samples were sent to all 3,490 men. The baseline examinations were carried out in 1973–1975, mainly in 1974; 1,815 men were found to be healthy, without diabetes, clinical CVD or regular medications.

Table 2. Baseline examinations in 1974.

<table>
<thead>
<tr>
<th>Examination</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical examination</td>
<td>Blood pressure, pulse, weight and height measurement,</td>
</tr>
<tr>
<td>Laboratory tests</td>
<td>Cholesterol, triglycerides, glucose, one-hour glucose and ECG</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>Past and current medications</td>
</tr>
<tr>
<td></td>
<td>Smoking and alcohol consumption</td>
</tr>
<tr>
<td></td>
<td>weight at 25 years of age,</td>
</tr>
<tr>
<td></td>
<td>Self-rated health on a five-step scale: “very good”, “good”, “fair”, “poor”, or “very poor”</td>
</tr>
<tr>
<td></td>
<td>Self-rated fitness on a five-step scale: “very good”, “good”, “fair”, “poor”, or “very poor”</td>
</tr>
<tr>
<td></td>
<td>Detailed information about PA in the past year</td>
</tr>
</tbody>
</table>

BMI was calculated as weight (kg) divided by height (m) squared. Midlife weight gain was calculated as weight in 1974 minus reported weight at 25 years of age.

To reflect the CVD risk at baseline, the Keys’ risk score was determined in 1974. The Keys’ risk score is calculated by multiple logistic equation of five characteristics including age, BMI, smoking, cholesterol and systolic blood pressure at baseline; the score indicates the risk of cardiac death or Q wave infarction per 5 years in 1,000 men (Keys et al. 1972).
4.4.1 Physical activity assessment in 1974

The questionnaire on PA was completed by 782 men out of 1,815 healthy men. The mean age of these 782 men was 47.8 years (SD 4.1) in 1974.

The men were also asked to give a global description of their leisure-time PA during the past year on a 4-step scale:

1. Reading, watching television, going to the cinema or other sedentary activity.
2. Walking, cycling, skiing, gardening, bowling, fishing or other light exercise weekly.
3. Jogging, running, skiing, swimming, tennis, badminton, heavy gardening or similar exercise weekly on a regular basis.
4. Regular vigorous/competitive exercise several times a week on a regular basis.

Because only 11 men reported competitive PA, groups 3 and 4 were combined in the analyses. Men answering yes to question 1 were categorized as low PA; 2 as moderate; and 3 and 4 as high PA group.

Preliminary analyses showed that global description of PA differentiated the mortality risk most consistently, and therefore more detailed analyses were performed comparing these groups.

4.5 Physical activity surveys in 1986 and 2000

In the 11-year survey in 1986, the men were asked two questions about PA:

1. Do you exercise regularly weekly?
2. If yes, how many hours per week?

Of the 782 men, 80.7% (n = 631) could be re-assessed with questionnaires in 1985–1986.

In the 26-year survey in 2000, men were asked three questions about their PA:

1. Do you exercise regularly weekly?
2. If yes, how many hours per week?
3. How many times a week do you have exercise leading to sweating and breathlessness?
Of the 782 men, 70.6% (n = 552) could be re-assessed with questionnaires in 2000.

For the 514 men included in the frailty analyses we evaluated the consistence of PA between 1974 and 2000. According to the reported hours used to exercise weekly (based on the first and second question described above), the men were categorized into three PA levels similar to the categorization in 1974: low PA including men with less than 2 hours of exercise per week; moderate PA including men with 2–6 hours of exercise per week, and high PA including men with more than six hours/week. Thereafter, five groups according to the change in PA between 1974 and 2000 were formed: constantly low, constantly moderate, constantly high, increase in category and decrease in category of PA.

4.6 Questionnaire surveys in 2000 for HRQL and frailty status

In 2000, after 26 years of follow-up, questionnaires were sent to the survivors, and re-sent once to the non-respondents. 552 men (mean age 73 years, range 66–81 years) responded, representing 91% of the survivors at that time and 70.6% of the original PA population.

The questionnaire included the following components:

1. Questions about PA as described above
2. The Finnish version of the RAND-36-Item Health Survey 1.0
3. Current diseases including CVD and diabetes

From the responses to the questions about current diseases, we calculated the Charlson index (Charlson et al. 1987), which is a co-morbidity index developed to control the effect of comorbid diseases and conditions in longitudinal studies.

4.6.1 HRQL assessment

HRQL was appraised from the answers to the 36 items (questions) of the RAND-36 questionnaire (Ware & Gandek 1998). The items of RAND-36 make up the eight scales which represent different HRQL dimensions, each of them consisting of 2–9 items. The scoring of the scales was conducted in two parts. First, the answers to each item were transformed to a 0 to 100 scale, so that the lowest and highest level of function and well-being were set at 0 and 100, respectively. Next, the scores to the scales were calculated by summing the scores of items in the
scale and dividing it by the number of items in the particular scale. The scores were calculated for all the eight domains.

4.6.2 Frailty assessment

For frailty assessment we used four criteria modified from Fried’s Scale (Sirola et al. 2011):

1. Shrinking; defined as weight loss of > 5% from baseline in 1974 or having current BMI below 21 kg/m2
2. Physical weakness; based on self-reported difficulty (not at all = 0) in carrying or lifting a grocery bag (one question of the physical function scale of RAND-36)
3. Exhaustion; based on reported low energy most or all of the time during the preceding 4 weeks (one question of the vitality scale of RAND-36)
4. PA; based on the response to the question: “Do you exercise regularly weekly?” The answer “No” was taken to denote low PA or sedentary living

The participant was classified as frail or prefrail if three to four, or one to two criteria were met, respectively, and nonfrail if none of the criteria were met. The frailty status was appraised only if answers to all four questions were provided.

4.7 Telomere length measurement

In 2002–2003, after 29-year follow-up, a random subcohort of all survivors (n=1,234) of the Helsinki Businessmen Study was invited to laboratory tests. 204 men (mean age 76 years) of this random subcohort were included in our cohort with PA assessment in 1974. Venous blood samples were taken for genetic analyses.

In addition to mean LTL, the proportion of short telomeres was of interest, as it has been suggested that one critically short telomere may lead to cellular senescence (Abdallah et al. 2009, Bendix et al. 2010, Hemann et al. 2001). As there is no exact knowledge as to how many telomere repeats are necessary to protect chromosomes, there are no universal recommendations for the threshold for very short telomeres (Bendix et al. 2010). In our study, the LTL of 5 kB was chosen for the threshold, as it was the lowest cut-off limit providing reliable results in preliminary analyses.
The mean LTL measurements and the proportion of short telomeres (<5 kB) were performed in the Minerva Institute for Medical Research in Helsinki, Finland. The extraction of DNA from peripheral blood leukocytes was performed with standard procedures using the PureGene, Gentra method (Gentra systems, Minneapolis, MN, USA). The mean LTL and the proportion of short telomeres (<5 kB) were measured using the Southern blot technique with TeloTAGGG Telomere length assay kits (Roche Molecular Biochemicals, Basel Switzerland). The procedures have been described previously (Fyhrquist et al. 2010).

4.8 Mortality follow-up

Total mortality of the study cohort from 1 January 1974 up to 31 March 2007 was retrieved from the National Population Information System of the Finnish Population Register Centre (Finnish Population Register Centre 2014), which keeps registry of all Finnish citizens. According to the Centre, assessment of vital status is very reliable for people having their permanent place of residence in Finland (over 95% of the study population) irrespective of whether they die in Finland or aboard. Furthermore, the assessment of vital status is also quite reliable for Finnish citizens living permanently aboard.

Causes of death up to 31 December 2006 were determined from the nationwide computerized Cause of Death Register of Statistics Finland, in which trained nosologists code the causes of death. The causes were categorized into 5 groups: cardiovascular, cancer, accidents, suicides and other causes, referring to the International Classification of Diseases codes I10-I99, C00-D48, W00-X59, X60-X87 and other codes, respectively.

4.9 Statistical methods

Continuous variables were compared across PA groups using T-test, nonparametric tests and analyses of covariance (ANCOVA), and categorical variables were compared using chi-square and trend tests (I-IV). Differences in survival curves were analysed with log rank test, and hazard ratios with their 95% confidence intervals (CI) for mortality associated with PA at baseline were calculated using Cox’s proportional hazards regression. Survival time was calculated as the number of years between baseline in 1974 and death or end of the follow-up in March 2007, whichever happened first (IV). Odd ratios with their 95% CIs for frailty or prefrailty associated with PA were calculated using
multinominal logistic regression models with nonfrail men as referents (II). Spearman rank correlation and analysis of covariance (ANCOVA) were used to study the relationship between PA, mean LTL and the proportion of short telomeres, and Fisher’s multiple-comparison test was used to compare groups (III).

Potential confounders (age, smoking, BMI, cholesterol, blood pressure, one-hour glucose and alcohol consumption in 1974, Keys’ risk index in 1974 and Charlson comorbidity index in 2000) in various combinations were adjusted for in the respective models. In statistical analyses 2-tailed tests were used and P values <0.05 were taken as significant (I-IV).

The analyses in all studies (I-IV) were performed using NCSS statistical software 2004 or 2007 version (NCSS, Kaysville, Utah, USA).
5 Results

As a global description of leisure-time PA 148 (18.9%), 398 (50.9%) and 236 (30.2%) of the men reported low (sedentary), moderate, and high PA, respectively. Most men (80.5%, n = 630) had sedentary work and commuted to work by car or public transportation. There was no association between PA groups and the type of work.

The most popular leisure-time activities among the 782 participants were gardening, wood chopping and home repairing and the most popular types of exercise were walking, skiing and swimming. Of the various types of exercise, only the frequencies of skiing, jogging and tennis were significantly different between groups (data not shown). As the global description of leisure-time PA most consistently differentiated the endpoints in the preliminary analyses, all the analyses presented in this thesis are performed comparing low, moderate and high PA groups.

The baseline characteristics of the study population according to the PA levels are presented in Table 3. Men in the high PA group tended to gain less weight during adulthood, have lower BMI in midlife and lower 1-h postload glucose concentration. Smoking was more common among sedentary men, but alcohol consumption was highest among men in the high PA group. Serum lipid levels and blood pressure were only weakly associated with PA, blood pressure being lowest among men in the moderate PA group.

PA was measured three times: at baseline and in the 11-year and 26-year surveys; however, the questions on PA were less detailed in the follow-ups.

The continuity of initially reported regular weekly exercise was significantly higher among the most physically active men in both subsequent surveys (P<0.001 and <0.002, respectively). “Vigorous” exercise leading to sweating and breathlessness was reported by 56, 68 and 79% of the men in the low, moderate and high PA groups in the 26-year survey, respectively, while 35, 45 and 60% performed vigorous exercise more often than twice a week as shown in Table 4.
Table 3. Characteristics of the study population in 1974 (IV, published by permission of Springer).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Whole group (n=782)</th>
<th>Low PA (n=148)</th>
<th>Moderate PA (n=398)</th>
<th>High PA (n=236)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age, year</strong></td>
<td>47.9 (4.1)</td>
<td>48.5 (3.9)</td>
<td>47.9 (4.2)</td>
<td>47.4 (4.1)</td>
</tr>
<tr>
<td><strong>BMI at 25 years of age, kg/m²</strong></td>
<td>22.7 (2.2)</td>
<td>22.5 (2.1)</td>
<td>22.7 (2.3)</td>
<td>22.9 (2.2)</td>
</tr>
<tr>
<td><strong>BMI in 1974, kg/m²</strong></td>
<td>26.3 (2.8)</td>
<td>26.5 (2.9)</td>
<td>26.6 (2.9)</td>
<td>25.6 (2.6)</td>
</tr>
<tr>
<td><strong>Weight gain from 25 years of age until 1974, kg</strong></td>
<td>11.2 (8.4)</td>
<td>12.5 (8.8)</td>
<td>12.3 (8.8)</td>
<td>8.7 (6.6)</td>
</tr>
<tr>
<td><strong>Perceived health, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very good or good</td>
<td>410 (53.9)</td>
<td>69 (48.6)</td>
<td>196 (50.5)</td>
<td>145 (63.0)</td>
</tr>
<tr>
<td>Fair</td>
<td>320 (42.1)</td>
<td>69 (48.6)</td>
<td>175 (45.1)</td>
<td>76 (33.0)</td>
</tr>
<tr>
<td>Poor or very poor</td>
<td>30 (3.9)</td>
<td>4 (2.8)</td>
<td>17 (4.4)</td>
<td>9 (3.9)</td>
</tr>
<tr>
<td><strong>Perceived fitness, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very good or good</td>
<td>231 (30.3)</td>
<td>24 (16.9)</td>
<td>86 (22.1)</td>
<td>121 (52.6)</td>
</tr>
<tr>
<td>Fair</td>
<td>420 (55.2)</td>
<td>77 (54.2)</td>
<td>243 (62.5)</td>
<td>100 (43.3)</td>
</tr>
<tr>
<td>Poor or very poor</td>
<td>110 (14.5)</td>
<td>91 (65.9)</td>
<td>60 (15.5)</td>
<td>9 (3.9)</td>
</tr>
<tr>
<td><strong>Keys’ risk index, %/5 years</strong></td>
<td>2.3 (1.9)</td>
<td>2.5 (1.9)</td>
<td>2.3 (2.1)</td>
<td>2.1 (1.6)</td>
</tr>
<tr>
<td><strong>Blood pressure, mmHg</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systolic</td>
<td>145.7 (18.8)</td>
<td>145.9 (17.4)</td>
<td>144.8 (19.2)</td>
<td>147.0 (18.2)</td>
</tr>
<tr>
<td>Diastolic</td>
<td>94.1 (10.7)</td>
<td>94.3 (10.1)</td>
<td>93.9 (10.9)</td>
<td>94.2 (10.8)</td>
</tr>
<tr>
<td>Pulse rate, per min</td>
<td>64.6 (10.8)</td>
<td>67.6 (11.8)</td>
<td>64.8 (10.3)</td>
<td>62.2 (10.5)</td>
</tr>
<tr>
<td>Cholesterol, mmol/l</td>
<td>6.4 (1.0)</td>
<td>6.6 (1.0)</td>
<td>6.4 (1.1)</td>
<td>6.4 (1.0)</td>
</tr>
<tr>
<td>Triglycerides, mmol/l</td>
<td>1.7 (0.9)</td>
<td>1.9 (0.9)</td>
<td>1.8 (0.9)</td>
<td>1.6 (0.8)</td>
</tr>
<tr>
<td>One-hour glucose, mmol/l</td>
<td>7.3 (2.2)</td>
<td>7.7 (2.1)</td>
<td>7.4 (2.3)</td>
<td>6.8 (2.0)</td>
</tr>
<tr>
<td>Smokers (%)</td>
<td>275 (35.2)</td>
<td>66 (44.6)</td>
<td>132.0 (33.7)</td>
<td>77 (32.6)</td>
</tr>
<tr>
<td>Alcohol, g/week</td>
<td>126.0 (56.0–252.0)</td>
<td>126.0 (42.0–252.0)</td>
<td>126.0 (59.5–248.5)</td>
<td>126.0 (63.0–280.0)</td>
</tr>
</tbody>
</table>

Median (IQ)

Units in parentheses: Mean (SD) unless otherwise stated.
Table 4. Reported PA in 1974 in relation to PA in the 11- and 26-year surveys (IV, published by permission of Springer).

<table>
<thead>
<tr>
<th></th>
<th>Low PA in 1974 (n = 148)</th>
<th>Moderate PA in 1974 (n= 398)</th>
<th>High PA in 1974 (n = 236)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weekly exercise in the 11-year survey, n (%)</td>
<td>59 (49.6)</td>
<td>220 (69.2)</td>
<td>176 (91.7)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Weekly exercise in the 26-year survey, n (%)</td>
<td>64 (70.0)</td>
<td>216 (80.0)</td>
<td>157 (85.3)</td>
<td>0.002</td>
</tr>
<tr>
<td>“Vigorous” exercise * in the 26-year survey, n (%)</td>
<td>51 (56.0)</td>
<td>184 (68.1)</td>
<td>146 (79.3)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>“Vigorous” exercise * ≥3 times/week in the 26-year survey</td>
<td>18 (35.3)</td>
<td>82 (44.6)</td>
<td>88 (60.3)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

*Reported exercise intensity leading to sweating and breathlessness

5.1 HRQL and physical activity

The HRQL assessment was completed for 522 respondents in the 2000 survey, representing 91% of the survivors at that time. In 2000, the men with low PA in midlife reported significantly higher prevalences of CVD ($P=0.02$), cerebrovascular disorders ($P=0.05$) and chronic obstructive pulmonary diseases ($P=0.04$).

After adjusting for age, CVD risk factors (BMI, glucose, smoking, cholesterol and blood pressure) and self-rated health at baseline, only physical function of the eight domains of HRQL assessed in old age was significantly related to PA in midlife. Average physical function scores in the low, moderate and high PA groups were 72, 75 and 80, respectively ($P=0.004$). Adjustment for CVD, cerebrovascular disorders and chronic obstructive pulmonary diseases and Charlson index attenuated the association but did not eliminate the significance ($P=0.01$ as diseases were included in the model; $P=0.02$ as the Charlson comorbidity index was included).

Those who reported high PA at baseline had higher scores also on physical role, general health, vitality and social functioning domains of the RAND-36 compared to those with low or moderate PA, but the differences did not reach statistical significance as shown in Figure 3.
5.2 Frailty and physical activity

Frailty status was assessed for 514 participants, representing 83% of the survivors in the year 2000. A total of 48 participants were classified as frail, representing 9.3% of all the 514 participants, and the proportion of frail men was inversely and significantly \( (P=0.001) \) related to PA level in midlife.

The PA level in midlife significantly predicted frailty and prefrailty in 2000, as shown in Table 5. In the high PA group, the risk of frailty was 80% lower compared to the low PA group after adjustment for age, BMI, smoking, blood
pressure, cholesterol and alcohol consumption in 1974 (Model B in Table 5), and 77% lower after including in the model the comorbid conditions in 2000 as reflected by the Charlson comorbidity Index (Model C in Table 5). The risk for pre-frailty was 47% lower compared to the low PA group after adjustment as in Model B and 45% lower after adjustment as in Model C (Table 5).

Table 5. Odds ratios of frailty in 2000 by the level of PA in 1974 using multinominal logistic regression models (II, published by permission of Oxford University Press).

<table>
<thead>
<tr>
<th>Stage of frailty in 2000</th>
<th>Low PA in 1974 n=87 (16.9%)</th>
<th>Moderate PA in 1974 n=256 (49.7%)</th>
<th>High PA in 1974 n=171 (33.3%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonfrail men as referent</td>
<td>Odd Ratios 95% CI</td>
<td>Odd Ratios 95% CI</td>
<td>Odd Ratios 95% CI</td>
</tr>
<tr>
<td>Prefrailty</td>
<td>Model A</td>
<td>1.0 (Referent)</td>
<td>0.92 (0.53–1.60)</td>
</tr>
<tr>
<td></td>
<td>Model B</td>
<td>1.0</td>
<td>0.98 (0.55–1.72)</td>
</tr>
<tr>
<td></td>
<td>Model C</td>
<td>1.0</td>
<td>1.00 (0.57–1.78)</td>
</tr>
<tr>
<td>Frailty</td>
<td>Model A</td>
<td>1.0</td>
<td>0.60 (0.27–1.33)</td>
</tr>
<tr>
<td></td>
<td>Model B</td>
<td>1.0</td>
<td>0.68 (0.30–1.54)</td>
</tr>
<tr>
<td></td>
<td>Model C</td>
<td>1.0</td>
<td>0.75 (0.32–1.74)</td>
</tr>
</tbody>
</table>

Model A: adjusted for age; Model B: adjusted for age, BMI, smoking, blood pressure, cholesterol and alcohol consumption in 1974; Model C: adjusted for B and for the comorbidity index in 2000.

We also tested the effect of baseline one-hour postload glucose as a covariate, but the relationship between PA and old age frailty was virtually unaltered (data not shown).

The comparisons of the relationship between the consistency of PA (between 1974 and 2000) and stage of frailty at old age showed that a decrease in PA was associated with higher and an increase in PA with lower prevalence of frailty (Table 6). None of the men with constantly high PA were frail in 2000 ($P = 0.001$).

As PA is also one of the four frailty criteria used in our study, we made an analysis by adjusting for reported exercise in 2000 with the purpose of investigating if PA in midlife simply predicts exercise at old age without a broader association to old age frailty. Even in these analyses, fully adjusted odds ratio of frailty in the high PA group was as low as 0.31 (95% CI 0.10–0.97), referring to the importance of midlife PA in the prediction of frailty at old age.

<table>
<thead>
<tr>
<th>Consistency of PA between 1974 and 2000</th>
<th>Stage of frailty in 2000</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nonfrail</td>
</tr>
<tr>
<td>Constantly</td>
<td>n (%)</td>
</tr>
<tr>
<td>Low, n = 27</td>
<td>1 (3.7)</td>
</tr>
<tr>
<td>Moderate, n = 139</td>
<td>66 (47.5)</td>
</tr>
<tr>
<td>High, n = 55</td>
<td>32 (58.2)</td>
</tr>
<tr>
<td>Increase, n=118</td>
<td>45 (38.1)</td>
</tr>
<tr>
<td>Decrease, n=175</td>
<td>57 (32.6)</td>
</tr>
</tbody>
</table>

* Global P value between groups: <0.001

5.3 Telomere length and physical activity

In the subcohort of 204 men with LTL measurements, 38 (18.6%), 95 (46.6%) and 71 (34.8%) men belonged to the low, moderate and high PA groups, respectively. The subcohort represented quite well the whole study population as shown in Table 7.

Table 7. Characteristics of the whole cohort and the subcohort of 204 men (III, published by permission of Elsevier).

<table>
<thead>
<tr>
<th>Variable in 1974</th>
<th>Whole cohort (n=782)</th>
<th>Subcohort (n=204)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>47.9 (4.1)</td>
<td>47.0 (4.0)</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>26.3 (2.8)</td>
<td>25.8 (2.8)</td>
</tr>
<tr>
<td>Smokers, percentages</td>
<td>35.2</td>
<td>31.4</td>
</tr>
<tr>
<td>Cholesterol, mmol/l</td>
<td>6.4 (1.0)</td>
<td>6.5 (1.0)</td>
</tr>
<tr>
<td>1h glucose, mmol/l median (IQ)</td>
<td>7.3 (5.0–8.0)</td>
<td>6.8 (5.4–8.7)</td>
</tr>
<tr>
<td>Alcohol, g/week, median (IQ)</td>
<td>126 (56–252)</td>
<td>126 (56–252)</td>
</tr>
<tr>
<td>Blood pressure, mmHg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systolic</td>
<td>145.7 (18.7)</td>
<td>144.3 (18.3)</td>
</tr>
<tr>
<td>Diastolic</td>
<td>94.1 (10.7)</td>
<td>93.3 (10.8)</td>
</tr>
</tbody>
</table>

Units in parentheses: Mean (SD) unless otherwise stated

After adjusting for age, BMI, cholesterol and smoking in 1974, the mean LTLs were 8.10 kB (standard error [SE 0.07]), 8.27 kB (SE 0.05) and 8.10 kB (SE 0.05) in the low, moderate and high PA groups, respectively ($P=0.03$). As illustrated in Figure 4, the mean LTL was significantly longer in the moderate than in the low or high PA groups. The proportion of short telomeres was significantly lower in
the moderate (11.35%, SE 0.25) than in the high PA group (12.39%, SE 0.29) in the adjusted analysis ($P=0.02$), but there was no significant difference between the proportion of short telomeres in the moderate and low PA groups (12.21%, SE 0.39) as illustrated in the Figure 4.

![Figure 4. LTL and the proportion of short telomeres according to the PA level. Bars represent means with SE. *Adjusted for age. **Adjusted for age, BMI, cholesterol and smoking in 1974 (III, published by permission of Elsevier).](image)

**5.4 Mortality and physical activity**

During the 34-year follow-up, 295 men died (37.8%), with the cause of death available for 282 deaths (95.6% of total). The main causes of death were cancers (101 deaths during follow-up) and CVD (99 deaths during the follow-up). Cause of death was undetermined in 13 men. Mortality was not significantly associated with the type of work (sedentary, mobile) or type of commuting to work (by car, by bus, walking or cycling).
Mortality per 1,000 person years was 16.5, 12.9 and 10.8 ($P = 0.01$) in the low, moderate and high PA groups, respectively. Unadjusted survival curves are shown in Figure 5 demonstrating the consistent difference between the groups.

![Survival curve, unadjusted](image)

**Fig. 5. Survival curve, unadjusted (IV, published by permission of Springer).**

Especially deaths caused by CVD were associated with PA in midlife: there were 57.7 (95% CI:37.0–85.7), 43.3 (95% CI:32.3–56.9) and 32.8 (95% CI:21.1–48.8) CVD deaths per 10,000 person years in the low, moderate and high PA groups ($P=0.033$), respectively.

Multivariate analyses (Table 8) indicate that the protective effect of PA was independent of traditional CVD risk factors assessed at baseline reflected by Keys’ risk index (age, BMI, cholesterol, smoking and systolic blood pressure). The results did not change after adjustment for 1-h glucose and triglycerides.
Table 8. Hazard ratios for total mortality according to PA groups during the 34-year follow-up (IV, published by permission of Springer).

<table>
<thead>
<tr>
<th>Model</th>
<th>Low PA in 1974 (n=148)</th>
<th>Moderate PA in 1974 (n=398)</th>
<th>High PA in 1974 (n=236)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hazard ratio (95% CI)</td>
<td>Hazard ratio (95% CI)</td>
<td>Hazard ratio (95% CI)</td>
</tr>
<tr>
<td>Model A</td>
<td>1.52 (1.10–2.11)</td>
<td>1.21 (0.92–1.59)</td>
<td>1.0 (Referent)</td>
</tr>
<tr>
<td>Model B</td>
<td>1.68 (1.18–2.39)</td>
<td>1.26 (0.93–1.69)</td>
<td>1.0</td>
</tr>
<tr>
<td>Model C</td>
<td>1.64 (1.15–2.34)</td>
<td>1.22 (0.90–1.64)</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Hazard ratio was calculated using the Cox proportional hazard model; Model A: adjusted for age; Model B: adjusted for Keys’ risk index, perceived health, and perceived physical fitness in 1974; Model C: adjusted for B and triglycerides and log one-hour glucose in 1974.
6 Discussion

As the quotation from Hippocrates in the Introduction illustrates, the first observations of the beneficial effects of PA on health and longevity were made a long time ago. During the last century, the impacts of PA and inactivity have been extensively investigated by improving research methods. The world’s population is ageing and strategies to enhance overall health, physical function and quality of live among older people are in great demand. The aim of this study was to investigate whether PA in midlife is associated to HRQL and frailty status in old age as well as to LTL, a potential indicator of cellular ageing, and mortality.

The present study is part of the Helsinki Businessmen Study, a cohort study gathered in 1964–1973 by the Helsinki Occupational Health Institute. The primary aim of the Helsinki Businessmen Study was to find healthy middle-aged men for a multifactorial primary prevention study.

As occupational health check-ups were not customary in the 1960s and 1970s, they were mostly performed for employees in leading positions. The participants were mostly business executives or had academic education and were from the highest social class.

This socioeconomic homogeneity of the study population is a unique characteristic of the study. Socioeconomic factors are known to be related to PA patterns. For example, leisure-time PA is shown to be more common in high social class and particularly in occupations with a high level of decision latitude (Kirk & Rhodes 2011), which probably illustrates well our population. High socioeconomic status has been suggested to be inversely associated to occupational PA, thus reducing the protective effect of occupational PA (Hu et al. 2004). Further, socioeconomic factors have been constantly associated to most of our outcomes like HRQL (Lahana et al. 2010, Robert et al. 2009), frailty (Alvarado et al. 2008, Cigolle et al. 2009, Lang et al. 2009a, Woo et al. 2005), and mortality (Doornbos & Kromhout 1990, Guralnik et al. 1993, Laaksonen et al. 2008, Lin et al. 2003, Liu et al. 1982, Salonen 1982, Sorlie & Rogot 1990). The findings about the relations between LTL and socioeconomic factors are controversial (Adams et al. 2007, Woo et al. 2008). Hence, the socioeconomical homogeneity may reduce confounding from this source.

In our cohort of 782 men, self-reported PA in midlife was associated with several important aspects of healthy ageing as well as longevity. The men in the high PA group reported significantly higher scores in the physical function domain of HRQL in old age and had a lower risk for frailty compared to the men...
in the low PA group. Moreover, none of the men who consistently reported high levels of PA were frail in 2000 when the mean age of participants was 74 years. In our analyses, men in the moderate PA group had longer LTL than the men in the low or high PA group, supporting the view that moderate levels of PA are beneficial in terms of LTL. Finally, there was a strong association between midlife PA level and mortality.

Although a cause-and-effect relationship cannot be drawn from observational studies like the present study, our findings support the view that those who have a physically active lifestyle in middle age may have better physical function in old age and lower risk for frailty and mortality.

### 6.1 HRQL and physical activity

We found that PA in midlife was associated with one of the eight domains of HRQL, i.e., physical function, which consists of ten questions in the RAND-36 questionnaire. The differences in physical function scores reached the MCID, minimal clinically important difference, which has been suggested to be 3–5 points for RAND-36 (Samsa et al. 1999). When comparing the lowest and highest level of PA, the difference was 8 points. In recent prospective studies, the measured differences in physical function associated with higher PA level have been approximately 3–5 points (Balboa-Castillo et al. 2011, Sanchez-Villegas et al. 2012, Oostrom et al. 2012). RAND-36 measures the participant’s own, subjective view of HRQL and physical function as part of it, but the physical function scores of RAND-36 have been shown to be a valid measure of mobility-disability (Syddall et al. 2009). Hence, the 8-point difference in physical function score may have a marked influence on the ability to live independently during the coming years and decades.

The adjustment for baseline self-rated health reduces the possibility of reverse causality, i.e., that the low level of PA in midlife is due to subjective poor health status (Bowling 2001). The causality caused by ageing, i.e., that older participants would be less active and have more diseases, is minimized by the use of baseline PA assessment in the analyses. In midlife the age differences do not affect PA engagement as much as in older age, and thus it is likely that age differences did not affect the PA categorization in our study. Furthermore, the analyses were adjusted for main diseases and the Charlson co-morbidity index in 2000, which did not significantly change the results.
Our findings are in line with previous cross-sectional studies, which have shown that PA is associated particularly with the physical function scale in RAND-36 (Franco et al. 2012, Wanderley et al. 2011). The fact that the differences in other domains did not reach statistical significance was contrary to our hypothesis that PA in midlife would be associated with a broader range of HRQL domains. In that respect our results are in contrast to recent longitudinal studies (Balboa-Castillo et al. 2011, Sanchez-Villegas et al. 2012, Oostrom et al. 2012, Wolin et al. 2007), which have observed positive associations between PA and both mental and physical domains of HRQL. Interestingly, in a study by Heesch and colleagues, the positive association with mental health attenuated during the 6-year follow-up (Heesch et al. 2012). As the associations between PA and HRQL have been stronger when using objective assessment of PA, such as accelerometers (Anokye et al. 2012), our subjective way to assess PA by a questionnaire may have diluted the associations.

Socioeconomic status has been shown to be an independent predictor of HRQL (Lahana et al. 2010, Robert et al. 2009) and also in Finland, low income and education are associated with poorer self-rated health (Mackenbach et al. 2008). In this respect, our socioeconomically homogenous cohort is a remarkable strength of this substudy. Furthermore, socioeconomical homogeneity may in part explain the different outcomes of our study in contrast to other studies on PA and HRQL.

### 6.2 Frailty and physical activity

PA in midlife was strongly associated to frailty in old age in our 26-year follow-up. The men in the high PA group had an even 80% lower risk for frailty compared with sedentary men even after adjusting for several potential confounding factors such as age, smoking, BMI, cholesterol, blood pressure and alcohol consumption assessed in midlife. When the comorbid conditions in 2000 were included in the model, the risk for frailty was still 77% lower in the high PA group compared to the low PA group, but the risk for prefrailty was not further statistically significantly different between groups. Furthermore, none of the men who reported high PA in 1974 and 2000 were frail in 2000. The prevalence of frailty in 2000 was explicitly highest among the men with consistently low PA during the follow-up.

As PA is both the independent variable (i.e., the exposure) and one of the four frailty criteria used in our study to form the outcome, it may raise the question of
whether the associations we observed between PA in midlife and frailty in old age demonstrate only that the physically active participants in mid-life are physically active also in old age. To enlighten this question, we adjusted the analysis for PA in old age; even in these analyses, the power of midlife PA to predict frailty prevailed. Furthermore, the long follow-up reduces the likelihood for that phenomenon. PA has been suggested to be one of the moderately important domains of frailty (Sourial et al. 2012) but frailty itself is a multifaceted syndrome (Abellan van Kan et al. 2008, Fried et al. 2004, Strandberg & Pitkala 2007).

Our findings suggest that PA in midlife and onwards has a unique long-term effect in preventing frailty in old age. The physical phenotype of frailty has been shown to predict several adverse outcomes, like falls, fractures, institutionalization and post-operative complications, and therefore, prevention of frailty is an important issue both from the individual and societal point of view. It has been discussed that forward-looking policy for advancing successful ageing should include programmes to promote healthy ageing from midlife on, rather than simply aiming to support elderly people with chronic conditions (Bowling & Dieppe 2005).

Although PA has been an object of increasing attention in frailty treatment and prevention, investigations of the long-term relationships between PA and frailty among initially healthy people or general populations are still scarce. To the best of our knowledge, there is only one previous longitudinal study examining the associations between long-term PA and old-age frailty (Peterson et al. 2009). In the 5-year follow-up of 2,018 participants, sedentary older adults were found to have increased odds of frailty compared with those who engaged regularly in exercise activities. Frailty was defined as having a gait speed of less than 0.6m/s or being unable to rise from a chair once with arms folded and it was determined at three time points: at baseline and after 3 and 5 years. In that study by Peterson and colleagues the baseline number of diagnoses was the strongest predictor of frailty. The mean age of the study population was 74 years, and they had various co-morbidities at baseline, such as hypertension, CVD and diabetes. (Peterson et al. 2009) In contrast, the participants in our study were middle-aged at baseline (mean age 48 years) and did not have regular medication or diagnoses of CVD or diabetes at baseline. Nevertheless, their finding that participation in self-selected structured exercise activities is associated with lower risk of old age frailty is in line with our results.
Various diseases, conditions and phenomena are connected to frailty, for example through aetiology and pathogenesis. PA has been shown to reduce the risk of many of them, such as sarcopenia (Peterson et al. 2010), cardiovascular and other diseases (Cornelissen & Fagard 2005, Jeon et al. 2007, Warburton et al. 2006), as well as cognitive decline (Sofi et al. 2011). Furthermore, better fitness in midlife has been associated to lower incidence of chronic conditions (Willis et al. 2012). These associations highlight the possible pathways through which PA can prevent frailty.

In both observational and intervention studies PA after 65 years of age has also been found to have a strong protective effect against mobility disability (Fries et al 1994, Paterson & Warburton 2010) and falls (Sherrington et al. 2008) in later life. In an observational study of decedent population, constant long-term PA since midlife was associated to less need for hospital care in men and for long-term care in women in the last year of life (von Bonsdorff et al. 2009). Our finding that PA in midlife is associated to lower risk of frailty in old age is in line with these previous studies suggesting that midlife PA is associated to better physical function in old age and less need of care at the end of life.

The prevalence of physical phenotype frailty was 9.3% in our socioeconomically homogeneous male cohort, which is quite similar to the findings of a recent review, in which the prevalence of physical frailty among community-dwelling older adults was approximately 9.9% (Collard et al. 2012). Nevertheless, social factors have been shown to affect the frailty syndrome (Alvarado et al. 2008, Cigolle et al. 2009, Lang et al. 2009a, Woo et al. 2005). In a cross-sectional study of 10,661 older adults in Latin America, gender, early-life conditions, lack of schooling, a manual occupation and perceived economic hardship were associated to physical phenotype frailty (Alvarado et al. 2008). Thus the socioeconomic homogeneity of our cohort may reduce confounding also in this substudy.

### 6.3 Telomere length and physical activity

In our data, long-term moderate level of PA was associated with longer mean LTL than both low and high levels of PA. After adjusting for age and possible confounding factors such as BMI, cholesterol and smoking in 1974, the difference between mean LTL of the moderate and low PA groups was 172 BP. As the annual mean LTL shortening rate has been suggested to be approximately 25 BP/year in cross-sectional studies and approximately 40 BP/year in a few
longitudinal studies (Muezzinler et al. 2013), the difference of 172 BP could be taken as a 5- to 7-year difference in the “biological age” of telomeres.

As it has been suggested that one critically short telomere may lead to cellular senescence regardless of mean LTL (Abdallah et al. 2009, Bendix et al. 2010, Hemann et al. 2001) the proportion of short telomeres was also determined. In line with our findings between PA level and LTL, the proportion of short telomeres was lower in the moderate PA group than in the high PA group.

Considering the robust evidence for the dose-response-like association between PA and mortality (Chakravarty et al. 2012, Moore et al. 2012, Samitz et al. 2011) as well as our previous finding of a strong association between PA and mortality among all 782 participants, who were reasonably well represented by the 204 men with LTL analyses, our finding that men in the high PA group had shorter mean LTL than men in the moderate PA group is surprising - and contrary to our hypothesis of finding the longest mean LTL in men in the high PA group. Our findings may be partly explained by the fact that LTL has been found to be a weaker predictor of mortality among older adults than among younger adults (Boonekamp et al. 2013, Cawthon et al. 2003, Harris et al. 2006, Martin-Ruiz et al. 2005). In a study by Cawthon and colleagues, longer LTL was not associated to better survival in adults above 74 years (Cawthon et al. 2003) and similar results have been found in a study with participants above 85 years (Martin-Ruiz et al. 2005). At the point of LTL measurements, the mean age of our study population was 76 years.

Similarly with our results, Ludlow and colleagues found an inverse u-shaped correlation between PA and LTL among 69 participants aged 50-70 years (Ludlow et al. 2008). Furthermore, athletes with exercise-associated fatigue, aged 22 to 59 years, were reported to have shorter muscle TL than their controls in a relatively small study of 26 participants (Collins et al. 2003). This inverted U phenomenon has been suggested to be possibly age-dependent and apparent only in older, highly active individuals (Ludlow et al. 2013). The absence of additional benefit from higher than moderate intensity exercise has been reported in a large cross-sectional analysis of 7,813 women aged 40 to 70 years. The less active women had shortest LTL, but there was no additional increase in LTL for the most active women compared to the moderately active women (Du et al. 2012).

In contrast to our results, several studies have reported the longest LTL at the highest levels of PA (Cherkas et al. 2008, Denham et al. 2013, Kim et al. 2012, LaRocca et al. 2010). It is also notable that not all studies have found significant associations between PA and LTL (Cassidy et al. 2010, Song et al. 2010b, Woo et
Currently, the relationship between PA and LTL is not fully established, particularly at higher levels of PA.

LTL has been inversely associated to inflammation in epidemiologic studies (Bekaert et al. 2007, O'Donovan et al. 2011) and oxidative damage has been found to be a major cause of telomere shortening in experimental studies (Kawanishi & Oikawa 2004, Kurz et al. 2004, Tchirkov & Lansdorp 2003). As regular PA has been associated to lower levels of inflammatory markers (Yu et al. 2009, Taaffe et al. 2000) as well as to a long-term increase in endogenous antioxidant enzyme activity (Fatouros et al. 2004, Michailid et al. 2007, Radak et al. 2008), oxidative stress and inflammation may be the key mechanisms explaining the associations between regular PA and LTL attrition. Interestingly, very low and very high intensity exercise have been proposed to have less beneficial effects on oxidative and inflammation systems than regular moderate intensity exercise (Michailid et al. 2007, Radak et al. 2008, Tsai et al. 2001), which could explain our results.

6.4 Mortality and physical activity

In our cohort of 782 men, self-reported PA was strongly associated to 34-year mortality independently of traditional CVD risk factors at baseline. Comparison of the causes of deaths demonstrated that especially deaths caused by CVDs were associated to PA level in midlife.

Our findings are in line with the previous cohort studies showing an inverse relationship between PA and mortality (Andersen et al. 2000, Chakravarty et al. 2012, Leon et al. 1987, Manini et al. 2006, Moore et al. 2012, Paffenbarger et al. 1986,) as well as with a recent meta-analysis showing a 29% lower risk for mortality in the highest PA groups compared to the lowest PA group after adjusting for possible confounding factors. (Samitz et al. 2011).

In our study, the mortality reduction in the moderately activity group did not reach statistical significance, but previous studies have shown a clear dose-response-like association between PA and mortality (Paffenbarger et al. 1986,
Wen et al. 2011) similarly with the recent dose-response meta-analyses (Janssen et al. 2013, Samitz et al. 2011). This demonstrates that in terms of longevity, increased levels of PA may be beneficial. In observational studies the dose-response-like association is also an important finding, which enhances the possibility of a causal relationship.

Significant differences in mortality rates became evident after approximately 15 years of surveillance, and were shown to increase with time over the 34-year follow-up. In the same way, a Swedish follow-up study showed that approximately 10 years of increased PA after midlife is required before the benefits of PA are seen in mortality rates (Byberg et al. 2009). These findings are interesting, because in a review of 41 studies on leisure-time PA, reductions in mortality tended to be smaller with a follow-up longer than 11 years than with a shorter follow-up (Samitz et al. 2011).

The association between higher levels of PA and lower mortality may be partly explained by lower prevalence of several diseases and conditions which cause excessive mortality but can be prevented by PA (Cornelissen & Fagard 2005, Haskell et al. 2007, Jeon et al. 2007, Lee et al. 2012, Warburton et al. 2006). Furthermore, higher levels of PA are associated with better cardiorespiratory fitness, which is further associated with better survival (Lee et al. 2011, McAuley et al. 2012, Myers et al. 2002). Nevertheless, the mechanisms by which regular PA promotes longevity are not fully explained.

Although PA level was inversely associated with weight gain during adulthood, BMI at baseline did not attenuate the protective effect of PA on mortality in our data. Similarly, in a large pooled analysis, the positive association between PA and longer life expectancy was found in all BMI categories, yet having normal weight was advantageous (Moore et al. 2012). In a study of 14,345 men with high socioeconomic status, cardiorespiratory fitness was associated with a lower risk of mortality regardless of BMI (Lee et al. 2011). Furthermore, the obesity paradox, which refers to the phenomenon that obese patients survive longer in populations with a high prevalence of CVD (McAuley & Blair 2011, Wei et al. 1999), may be largely explained by cardiorespiratory fitness (McAuley & Beavers 2014). These findings are highly important to consider, as PA can easily be valued simply as a part of weight control. The baseline BMI, albeit together with other CVD risk factors, has also been included in other analyses of this thesis (HRQL, frailty, and LTL analyses) without substantial changes in outcomes. This further emphasizes that PA deserves to be promoted and accredited as an independent source of health, quality of life and longevity.
As socioeconomic status, as measured by income, education or occupation, has been found to be associated to mortality in several epidemiologic studies (Doornbos & Kromhout 1990, Guralnik et al. 1993, Laaksonen et al. 2008, Lin et al. 2003, Liu et al. 1982, Salonen 1982, Sorlie & Rogot 1990), the socioeconomic homogeneity of the study cohort may reduce confounding from socioeconomic factors.

6.5 Strengths and limitations

6.5.1 Strengths

The obvious strength of this study is the long follow-up time during which the participants were contacted several times. In the analyses, follow-up time ranged from 26 years to 34 years, which is exceptional in the context of studies focusing on the associations between PA and HRQL, frailty and TL.

High social class has been associated to higher response rates (Hemingway et al 1997), and better adherence (Heesch et al. 2012) in studies using RAND-36. The response rate, 91%, for the questionnaire survey in 2000 was relatively high, as the completion rates of RAND-36 are usually approximately 70% for postal surveys (Djarv et al. 2013, Jenkinson et al. 1994). For the evaluation of frailty, 83% of the participants answered the required four questions in 2000. Moreover, as social factors are related to PA patterns, as well as to most of our outcomes, the confoundings from that source may be reduced because of our socioeconomically homogeneous cohort.

6.5.2 Limitations of the physical activity assessment and categorization

In our study, PA was assessed using a self-reported questionnaire. As subjective methods, PAQs are prone to misclassification and overestimation of PA (Banda et al. 2010, Matton et al. 2007, Troiano et al. 2008, Troiano 2009), especially at the lower levels of PA (Aadahl et al. 2007, Emaus et al. 2010, Friedenreich et al. 2006, Jacobs et al. 1993). Furthermore, social desirability (Klesges et al. 2004, Sallis & Saelens 2000) may have caused bias. If there is more overestimation of PA at lower levels of PA than at other PA levels, it may lead to overestimation of the true effect of PA. Nevertheless, in general PAQs are suggested to lead to

The categorization of PA was based on the global description of leisure-time PA as described above. PA was first divided into four groups, but as only few (n=11) men reported vigorous activity, the two highest levels were combined. Although this categorization can be said to be crude, and there is evidently large variation in PA within each group, it is clear that the alternative global descriptions represent ordered levels of total volume of leisure-time PA. Moreover, PAQs have been suggested to be reasonably good particularly in classifying participants into groups of general absolute intensity categories, e.g. light, moderate and vigorous; difficulties have been suggested to increase when trying to identify special activities of daily living, especially light to moderate intensity activities (Haskell 2012). With this in mind, it is not surprising that the global description of PA differentiated the outcomes most consistently in the preliminary analyses, instead of the other questions of the PAQ. Despite this, our study shows an independent association between PA and reduced premature mortality in the long follow-up. Furthermore, our cohort includes men with high socioeconomic status, and self-reported PA has been shown to be a better proxy for fitness among highly educated persons (Gerrard 2012), referring to the more accurate self-reports in this group.

Another limitation is that the PAQs sent during the follow-up (in 1986 and 2000) were not uniform with the baseline questionnaire. Most of the analyses were carried out considering only baseline PA level. This does not take into account the change in PA level during the follow-up, but it may minimize the possible reverse causality with ageing, i.e., that a low level of PA is due to diseases and poor health status associated with older age. Reverse causality may occur also with physical function: lower physical function and capacity may lead to a decreased level of PA and strengthen the association between PA and frailty, which is linked to disability.

6.5.3 Limitations of the cohort and settings

Our cohort consists only of men, which restrains the generalization to women. In large reviews and meta-analyses, PA has been found to have even more protective effect against mortality among women than men (Nocon et al. 2008, Samitz et al. 2011). While the high social class of the entire cohort is one of the strengths of the study, it also constricts the generalization to other social groups.
As only 782 participants of the Helsinki Businessmen Study cohort reported details of their PA in 1974, the sample sizes, 782, 522, 514 and 204 in mortality, HRQL, frailty and telomere analyses, respectively, were relatively small. Nevertheless, sample sizes were sufficient to reach statistically significant results in several analyses, especially between the high and low levels of PA.

During this study the participants were clinically examined only at baseline, while the surveys in old age rely solely on self-reported data, except for the mortality and telomere analyses.

The definitions of HRQL and frailty status were based on the questionnaire data gathered in 2000, when the mean age of the participants was 73 years. The RAND-36 questionnaire is quite comprehensive and the completion of the questionnaire has been shown to be a demanding task for older people, especially for those with mental or physical disabilities (Hemingway et al. 1997, Parker et al. 2006). Thus, those who did not respond to the survey in 2000 or who failed to complete the questionnaire as required for the HRQL or frailty evaluation may have had lower HRQL or been more frail compared to the respondents, leading to conservative bias.

Weight as well as diseases and medical conditions in 2000 were also based on self-reports and are thus susceptible to recall bias. However, the recall of weight even over long periods has been shown to be highly accurate (Tamakoshi et al. 2003), and self-reports tend to be more correct for serious diseases such as myocardial infarction, cancer and diabetes (Bergmann et al. 2004).

The long duration of the follow-up makes HRQL and frailty analyses susceptible for selection bias through mortality. Frailty has been associated with higher mortality in our cohort (Sirola et al. 2011) and in other populations (Avila-Funes et al. 2008, Bandeen-Roche et al. 2006, Cawthon et al. 2007, Ensrud et al. 2008, Fried et al. 2001, Rockwood et al. 2007, Woods et al. 2005). Furthermore, low HRQL (Fan et al. 2002, Kroenke et al. 2008, Tsai et al. 2007) and low PA (Samitz et al. 2011) are linked to higher mortality. Hence, there has probably been selection through mortality and those who had died during the follow-up have possibly had lower HRQL and physical function.

For frailty assessment, we used four criteria modified from the Fried Scale. As walking speed was not measured in 2000, it was not included in our frailty assessment. Although walking speed has been found to have a strong predictive value even when used separately (Abellan van Kan et al. 2009b, Studenski et al. 2011), the four criteria we used for frailty have been shown to predict mortality, development of mobility disability and slow walking speed during a 7-year
follow-up in our cohort (Sirola et al. 2011). Moreover, the physical frailty assessment using data from the RAND-36 without walking speed has been shown to be strongly associated with the risk of death, disability, hip fracture and hospitalization (Woods et al. 2005). In our frailty assessment, the weight loss was based on long-term change during the entire 26-year follow-up. In the original Fried’s Scale the weight loss over the past one year is taken into account (Fried et al. 2001).

It is unlikely that any of the participants would have been frail at baseline, as they were actively working and only one man rated his health as “very poor” in 1974. Nevertheless, the frailty stage was assessed only in 2000 and the onset of frailty during the 26-year follow-up is unknown.

The LTL analysis in our study was performed by TRF analysis using Southern blot technique, which is widely used and considered as “gold standard” for TL measuring techniques, but has some disadvantages important to consider. The mean TL provided by the TRF analysis includes subtelomeric regions (Cawthon 2002, Counter et al. 1992, Steinert et al. 2004), suggested to vary between individuals by as much as 2 kB (Cawthon 2002). This may have caused bias, which is most probably nondifferential and may have diluted our observations.

6.6 Conclusions

1. PA in midlife was associated with better physical function in old age, but was not significantly associated with other domains of health-related quality of life. The relationship was not explained, although it was attenuated, by diseases associated with less PA in our cohort.
2. PA in midlife was strongly and independently associated with the physical phenotype of frailty in old age. Moreover, none of the men with constantly high PA were frail in old age. The association between high PA in midlife and lower risk for frailty in old age remained even after considering the chronic diseases and PA level in old age.
3. Moderate level of PA in midlife was associated with longer LTL in old age compared to low and high levels of PA. These results indicate that moderate level PA may be beneficial in terms of LTL, but they also reflect the multifaceted effects of PA on cell biology and underscore the need for caution when interpreting the associations between LTL and PA.
4. Self-reported PA in midlife was strongly associated with better long-term survival in our 34-year follow-up of initially healthy men. This association was independent of traditional CVD risk factors: age, cholesterol, triglycerides, smoking, systolic blood pressure, 1-hour glucose and BMI at baseline.

5. Finally, the overall results of this thesis lend support to the view that those who have adopted a physically active lifestyle in midlife have better physical function and lower prevalence of frailty in old age and lower mortality.
References


Original articles


Reprinted with permission from American Medical Association (I), Oxford University Press (II), Elsevier (III) and Springer (IV).

Original publications are not included in the electronic version of the dissertation.
1262. Roisko, Riikka (2014) Parental Communication Deviance as a risk factor for thought disorders and schizophrenia spectrum disorders in offspring: The Finnish Adoptive Family Study

1263. Åström, Pirjo (2014) Regulatory mechanisms mediating matrix metalloproteinase-8 effects in oral tissue repair and tongue cancer

1264. Haikola, Britta (2014) Oral health among Finns aged 60 years and older: edentulousness, fixed prostheses, dental infections detected from radiographs and their associating factors


1266. Kuusisto, Sanna (2014) Effects of heavy alcohol intake on lipoproteins, adiponectin and cardiovascular risk

1267. Kviniemi, Marjo (2014) Mortality, disability, psychiatric treatment and medication in first-onset schizophrenia in Finland: the register linkage study


1269. Väyrynen, Juha (2014) Immune cell infiltration and inflammatory biomarkers in colorectal cancer

1270. Silvola, Anna-Sofia (2014) Effect of treatment of severe malocclusion and related factors on oral health-related quality of life


1273. Häkli, Sanna (2014) Childhood hearing impairment in northern Finland: prevalence, aetiology and additional disabilities

1274. Lehto, Liisa (2014) Interactive two-step training and management strategy for improvement of the quality of point-of-care testing by nurses: implementation of the strategy in blood glucose measurement

1275. Harjunen, Vanessa (2014) Skin stem cells and tumor growth: Functions of collagen XVIII in hair follicle cycling and skin cancer, and Bmx tyrosine kinase in tumor angiogenesis
Salla Savela

PHYSICAL ACTIVITY IN MIDLIFE AND HEALTH-RELATED QUALITY OF LIFE, FRAILTY, TELOMERE LENGTH AND MORTALITY IN OLD AGE