A life-course approach to chronic kidney disease
To Anette,
Felix, Julia, Casper and Engla

"True wisdom comes to each of us when we realize how little we understand about life, ourselves, and the world around us."

Socrates 469-399 BC
Örebro Studies in Medicine 196

Per-Ola Sundin

A life-course approach to chronic kidney disease – risks and consequences
© Per-Ola Sundin, 2019

Title: A life-course approach to chronic kidney disease – risks and consequences
Publisher: Örebro University 2019
www.oru.se/publikationer-avhandlingar

Print: Örebro University, Repro 08/2019

ISSN 1652-4063
Abstract


Successful primary prevention of chronic kidney disease (CKD) relies on understanding the pathways leading to established disease, including how they extend over the life-course. Projects in this thesis examine risk factors for CKD and consequences of impaired kidney function from a life-course perspective using routinely collected health-data in Swedish registers and research cohort data from the United Kingdom.

The main findings regarding risk factors for CKD are, that markers of health and development determined at conscription assessment in adolescence, independently predict diagnosis of end-stage renal disease in middle age. We also identified a persistent increased risk of CKD following hospital admission with pneumonia in adulthood with highest magnitude risks in years immediately following infection, but still statistically significantly raised more than 15 years after the pneumonia episode. Our main findings relevant to predicting the consequences of impaired kidney function are that creatinine and cystatin C used clinically to estimate kidney function (estimated glomerular filtration rate, eGFR) have associations with increased mortality risk independent of GFR measured with an exogenous filtration marker (mGFR). If cystatin C and creatinine are combined, adding mGFR does not improve mortality risk prediction. Another important finding is that moderately reduced eGFR is only associated with a statistically significant increased mortality risk among individuals in the lowest third of the distribution of grip strength in a general population sample followed for 4-5 years, after adjustment for potential confounding factors.

These results highlight the importance of adopting a life-course perspective when studying risk factors for CKD, since these associations can extend over different stages in the life-course. When assessing increased mortality risk associated with measures of GFR, combining cystatin and creatinine improves risk prediction. Potential effect modification across subgroups, including by grip strength, should be considered.

Keywords: chronic kidney disease, pneumonia, grip strength, creatinine, cystatin C, adolescence, life-course epidemiology, risk factor, mortality

Per-Ola Sundin, School of Medical Sciences, Örebro University, SE-701 82 Örebro, Sweden, perola.sundin@regionorebrolan.se
# Table of Contents

LIST OF ORIGINAL PAPERS .................................................................................................................. 11

ABBREVIATIONS .................................................................................................................................. 12

INTRODUCTION ..................................................................................................................................... 13
Life-course epidemiology .................................................................................................................. 14
Definition of CKD ................................................................................................................................. 14
Identification of CKD ............................................................................................................................... 15
  Measured GFR .................................................................................................................................... 15
  Estimated GFR ................................................................................................................................... 16
  Creatinine .......................................................................................................................................... 16
  Cystatin C .......................................................................................................................................... 18
  Creatinine and cystatin C combined................................................................................................. 19
Prevalence of CKD .................................................................................................................................. 19
Risk factors for CKD ............................................................................................................................... 20
  Inflammation and CKD ..................................................................................................................... 21
Infections and CKD ................................................................................................................................. 21
Consequences of CKD ............................................................................................................................ 21
eGFR and associations with adverse outcomes ..................................................................................... 22
Sarcopenia .............................................................................................................................................. 24

AIMS ...................................................................................................................................................... 25

MATERIAL AND METHODS .................................................................................................................. 26
Data sources .......................................................................................................................................... 26
  National population and health registers ......................................................................................... 26
  *The Total Population Register* ........................................................................................................ 26
  *The Swedish Population and Household Censuses* ...................................................................... 26
  *The Cause of Death Register* ........................................................................................................ 26
  *The Swedish Prescribed Drug Register* ........................................................................................ 27
  *The National Patient Register* ....................................................................................................... 27
  *The Swedish Military Conscription Register* .............................................................................. 27
  The Conscription Cohort ................................................................................................................... 28
  The mGFR Cohort .............................................................................................................................. 28
  The United Kingdom Household Longitudinal Survey ..................................................................... 29
Statistical methods and concepts .......................................................................................................... 30
  Cox regression .................................................................................................................................. 30
  Time dependent exposures and time dependent effects ............................................................... 30
Logistic regression ................................................................. 31
Conditional logistic regression .............................................. 31
Multivariable fractional polynomial method ......................... 32
Incidence density sampling ................................................. 33
Net reclassification improvement ........................................ 33
Multicollinearity ................................................................. 34
Effect modification and interactions ..................................... 35
Individual study designs and methods ................................. 36
Paper I – Predictors in adolescence of ESRD in middle-aged men 36
Study design ................................................................. 36
Statistical analysis ............................................................ 37
Paper II – Measured glomerular filtration rate does not improve prediction of mortality by cystatin C and creatinine ............... 38
Study design ................................................................. 38
Statistical analysis ............................................................ 38
Paper III – Hospital admission with pneumonia and subsequent persistent risk of chronic kidney disease ...................... 42
Study design ................................................................. 42
Statistical analysis ............................................................ 42
Paper IV – Grip strength modifies the association between estimated glomerular filtration rate and all-cause mortality ........ 45
Study design ................................................................. 45
Statistical analysis ............................................................ 46
ETHICAL CONSIDERATIONS ............................................. 48
RESULTS ............................................................................ 49
Paper I – Predictors in adolescence of ESRD in middle-aged men 49
Paper II – Measured glomerular filtration rate does not improve prediction of mortality by cystatin C and creatinine ............... 50
Paper III – Hospital admission with pneumonia and subsequent persistent risk of chronic kidney disease ...................... 51
Paper IV – Grip strength modifies the association between estimated glomerular filtration rate and all-cause mortality ........ 54
DISCUSSION ...................................................................... 56
Findings and implications .................................................. 56
Risks for CKD from a life-course perspective ......................... 56
Consequences of CKD ....................................................... 58
The Shrunken Pore Syndrome ............................................. 59
List of original papers


4. Sundin PO, Udumyan R, Fall K, Montgomery S. Grip strength modifies the association between estimated glomerular filtration rate and all-cause mortality. Accepted for publication in Nephrol Dial Transplant.

Published papers have been reprinted with permission from the publisher.
Abbreviations

BMI Body mass index
BSA Body surface area
CI Confidence interval
CKD Chronic kidney disease
CVD Cardiovascular disease
eGFR Estimated glomerular filtration rate
ESR Erythrocyte sedimentation rate
ESRD End-stage renal disease
EVF Erythrocyte volume fraction
GFR Glomerular filtration rate
HR Hazard ratio
ICD International Statistical Classification of Diseases and Related Health Problems
KDIGO Kidney disease: improving global outcomes
mGFR Measured glomerular filtration rate
NPR National patient register
NRI Net reclassification improvement
OR Odds ratio
Introduction

Chronic kidney disease (CKD) defined by reduced glomerular filtration rate (GFR) and the presence of albuminuria is an increasing public health issue with an estimated global prevalence of 8-16% (1). The estimated global number of deaths attributable to CKD has increased substantially during the last decades, including a rise of 31.7% between 2005 and 2015 reaching an estimated 1.23 million deaths (2). This increase is largely explained by population ageing and by increased prevalence of diabetes and hypertension which are major risk factors for CKD (1-3).

The long preclinical latency of CKD provides an opportunity for early identification of affected individuals and secondary preventive strategies including control of blood pressure preferably by agents blocking the renin-angiotensin system, glycaemic control in diabetes mellitus and lipid lowering therapy to reduce progression of CKD and its consequences. Due to lack of randomized controlled trials there is no consensus on potential benefits and cost-effectiveness of screening in the general population. Screening for CKD is currently recommended only in high-risk populations including individuals with diabetes, hypertension and cardiovascular disease (CVD) (4, 5).

Successful primary prevention of CKD relies on understanding not only the pathways leading to established disease but also insights into which age-defined time-periods these pathways can originate and how they extend over the life-course. Projects in this thesis examine risk factors for CKD and consequences of impaired kidney function from a life-course perspective using routinely collected health-data in Swedish registers, data from mandatory conscription examinations in late adolescence and research cohort data from the United Kingdom (UK).

To identify individuals with CKD and to quantify the increased risks associated with reduced GFR, the choice of method to assess GFR is of major importance. Major international clinical guidelines recommend the use of equations based on the serum concentration of the endogenous filtration markers creatinine and/or cystatin C; age; sex; and ethnicity to estimate GFR (eGFR) (5, 6). However, factors other than GFR may influence the serum concentration of the filtration marker and these non-GFR-factors may potentially be associated with the outcome of interest. In that case, assessment of risk attributed to reduced GFR will be confounded by associations between the filtration marker and the outcome not explained by GFR.
Another group of projects in this thesis are concerned with prediction of mortality from endogenous filtration markers used to identify CKD. In a large clinical sample generated locally, possible associations of cystatin C and creatinine with increased all-cause mortality risk independent of GFR determined with an exogenous filtration marker (mGFR), and the potential benefit of combining these measures were evaluated. The main non-GFR determinant of the serum concentration of creatinine is the production rate of creatinine in muscle tissue (7). In another project, possible effect modification by grip strength, as a marker of muscle status, on the association between GFR estimated from serum creatinine and all-cause mortality was investigated in the large Understanding Society cohort, which is representative of the general population in the UK (8).

**Life-course epidemiology**

Projects in this thesis aim to study risk of CKD and consequences of CKD with a life-course approach. The concept of life-course epidemiology has been described as

> “the study of long term effects on later health or disease risk of physical or social exposures during gestation, childhood, adolescence, young adulthood and later adult life” (9).

The aim is to identify processes that operate across an individual’s life course, or across generations, to influence the development of disease risk in contrast to the adult lifestyle model of adult chronic disease that focuses on how adult behaviour (notably smoking, diet, exercise and alcohol consumption) affect the onset and progression of diseases in adulthood (10, 11). Life-course epidemiology does not deny the importance of conventional risk factors but seeks to incorporate how biological and social factors throughout life independently, cumulatively and interactively influence health in adult life including how earlier life factors contribute in conjunction with these later life conventional risk factors to identify the pattern of risk development across the life-course (10).

**Definition of CKD**

A current widely accepted definition and classification of CKD based on reduction of GFR, albuminuria and other signs of kidney damage, present for more than three months, is included in the KDIGO (Kidney Disease: Improving Global Outcomes) guidelines (table 1) (5). This definition relies
heavily on determination of GFR which potentially has major implications for CKD research.

<table>
<thead>
<tr>
<th>Stage of chronic kidney disease</th>
<th>GFR&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Normal or high GFR with signs of kidney damage&lt;sup&gt;b&lt;/sup&gt;</td>
<td>&gt;90</td>
</tr>
<tr>
<td>2. Mild reduction of GFR with signs of kidney damage&lt;sup&gt;b&lt;/sup&gt;</td>
<td>60-89</td>
</tr>
<tr>
<td>3a. Mild to moderate reduction of GFR</td>
<td>45-59</td>
</tr>
<tr>
<td>3b. Moderate to severe reduction of GFR</td>
<td>30-44</td>
</tr>
<tr>
<td>4. Severe reduction of GFR</td>
<td>15-29</td>
</tr>
<tr>
<td>5. Kidney failure</td>
<td>&lt;15</td>
</tr>
</tbody>
</table>

<sup>a</sup>Glomerular filtration rate in ml/min/1.73 m<sup>2</sup> BSA. <sup>b</sup>Albuminuria >30 mg/day, urine sediment abnormalities, abnormalities detected by histology or imaging, electrolyte abnormalities due to tubular dysfunction or a history of kidney transplant.

Each stage of CKD may be further classified according to the amount of albuminuria.

**Identification of CKD**
The GFR equals the volume of primary urine filtered from the plasma by the renal glomeruli each minute. This absolute GFR is normalized for body surface area (BSA) to produce the relative GFR which, being the most important measure of kidney function, is included in the definition of CKD.

Reliable determination of GFR is of fundamental importance in many clinical situations beyond identifying and staging of CKD, including the evaluation of kidney function in patients with renal diseases, predicting the risk of disease progression, monitoring changes in renal function over time, determining the need to initiate dialysis therapy, screening potential living kidney donors and enabling dose adjustment of drugs cleared by the kidneys including potentially nephrotoxic agents in patients with impaired renal function (12).

GFR cannot be measured directly and therefore needs to be assessed indirectly by the kinetics of substances filtered in the glomeruli. An optimal filtration marker is inert, freely filtered in the glomeruli, not bound to proteins, not metabolized by the kidney and neither reabsorbed nor secreted in the renal tubules. An optimal endogenous filtration marker is also produced at a constant rate.

**Measured GFR**
The most reliable assessment of GFR is produced by observing the excretion of a known amount of an exogenous filtration marker administered to the
patient intravenously – measured GFR (mGFR). The golden standard of measured GFR is urinary inulin clearance which requires the continuous intravenous administration of inulin to maintain a stable serum concentration and bladder catheterization for urinary collection (13). Since the description of urinary inulin clearance in 1951, several alternative less complex methods to measure GFR have been developed, including plasma clearance of inulin, plasma or urinary clearance of chromium 51-labeled ethylenediaminetetraacetic acid (51Cr-EDTA), iothalamate and iohexol (14).

Plasma clearance of iohexol is an alternative procedure which only requires an intravenous bolus injection of iohexol followed by repeated venous blood sampling. In an extensive systematic review, plasma clearance of iohexol was assessed as an accurate method comparable to urinary inulin clearance (14). The analytical variation and the considerable within-subject biological variation of mGFR determined by plasma clearance of iohexol have been quantified as a total coefficient of variation in the range of 5-11% (14, 15).

**Estimated GFR**

Measuring GFR by exogenous filtration markers is a cumbersome and expensive procedure applied in clinical practice only when exact determination of GFR is essential. Instead, GFR is usually estimated from endogenous filtration markers, mainly serum creatinine but also to an increasing extent from serum cystatin C.

The precision required in estimates of GFR depends on its application. In the 2002 KDOQI guidelines an eGFR within 30% of mGFR was deemed satisfactory for clinical interpretation. The recommended measure of accuracy was the proportion of estimates within 30% of the mGFR (P30), and adequate equations for eGFR should have a P30 of at least 90% in their validation population (16). For research purposes a higher accuracy would be valuable when evaluating outcomes using eGFR as the exposure and actually quite necessary to successfully evaluate GFR or decline in GFR as a primary outcome (12).

**Creatinine**

The serum concentration of creatinine, first proposed as a filtration marker in 1926, is the most widely used marker of renal function in routine clinical practice (17). An international standardization for creatinine analyses using isotope dilution mass spectrometry (IDMS) was introduced in 2006 (18).
The reciprocal of the creatinine concentration (creatinine\(^{-1}\)) is proportional to GFR. However, a wide range of mGFR may be represented by the same creatinine level, due to influence from non-GFR determinants of the creatinine concentration. Generation of creatinine increases with dietary intake of increasing amounts of meat and protein, while malnutrition is associated with reduced creatinine generation. Generation of creatinine in muscle tissue is proportional to total muscle mass. Muscle mass is related to ethnicity, and especially the larger muscle mass in African-Americans, which has implications for the creatinine kinetics. Creatinine is also secreted by tubular cells and this represents an increasing proportion of the total urinary clearance of creatinine with decreasing GFR, reaching over 50% in advanced CKD (7, 14). Thus, serum creatinine alone is not a satisfactory indicator of GFR.

The 24 hour urinary clearance of creatinine is considered a surrogate marker of GFR but requires collection of urine. Besides being impractical, the collection of urine can result in error. Therefore, efforts have been made to develop methods to estimate GFR from the serum concentration of creatinine without collection of urine. The first of attempt to estimate GFR from serum creatinine was published in 1957 (19). The widespread clinical application of equations to estimate GFR from serum creatinine came with the publication of Cockroft and Gault’s formula in 1976, which included body weight, age and sex as surrogates for the non-GFR factors affecting serum creatinine (20). This equation estimates creatinine clearance rather than mGFR which incorporates an overestimation of GFR due to tubular secretion, is not related to BSA and was not developed with creatinine measurements using the same standardization as in current laboratories. However, creatinine clearance is still frequently used in instructions for dosage of drugs cleared by the kidneys, and in this setting measures of GFR should not be standardized for BSA since it is the capacity to excrete a substance that is of interest, not kidney function relative to body size.

Further research has led to the publication of over 40 different formulae to estimate GFR from serum creatinine with increasing mathematical complexity (12). Major contributions with broad clinical implications were the development of the Modification of Diet in renal Disease (MDRD) equation in 1999, later modified in 2006 to incorporate creatinine values traceable to the current IDMS standard (18, 21, 22). The MDRD equation generated from a population with CKD, using age, sex and ethnicity as surrogates for the non-GFR determinants of serum creatinine was the standard equation for several years. In 2009 the equation currently recommended in KDIGO...
guidelines, the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation was introduced (5, 23). This equation was developed and validated in a population with a wide range of mGFR in order to address the systematic underestimation of mGFR and imprecision of the MDRD equation at higher levels of mGFR. Despite a complex equation conditioned on level of creatinine in addition to factors representing age, sex and ethnicity the percentage of estimates within ±30% of mGFR (P30) was no more than 84.1%. The P30 value of the CKD-EPI equation differs in different populations and has been described to range from 67% to 87% in a recent review (12). There are limited data on the proportion of estimates within ±10% (P10) but available evidence suggest a value in order of 35-40% (12, 24).

Cystatin C
Given the limitations of eGFR based on creatinine, several alternative endogenous filtration markers have been investigated, but only serum cystatin C has been implemented widely in clinical practice. This low molecular weight protein was first suggested to be a marker of GFR 1979 and proposed as an endogenous filtration marker in 1985 (25, 26). Widespread use of cystatin C to estimate GFR was initiated in 1994 by a paper from Grub et al. comparing the performance of serum creatinine and cystatin C in estimating GFR, and introducing a new automated particle-enhanced turbidimetric method to measure cystatin C (27). An international standardization for cystatin C analyses was established in 2010 (28, 29).

The physiological function of Cystatin C is to inhibit the activity of cysteine proteases central in both inflammation and degradation of damaged proteins (30). Cystatin C is produced at a nearly constant rate in all nucleated cells and is not influenced by dietary intake of meat or tubular secretion and only marginally by muscle mass. Cystatin C is completely metabolized by tubular cells and is therefore not present in the urine (31). Thus, cystatin C cannot be used to calculate urinary clearance. The half-life of cystatin C is shorter compared with creatinine which gives the advantage of earlier detection of acute changes in GFR (32).

As in the case of creatinine, a specific serum cystatin C value may represent a wide range of mGFR. Judging from the accuracy of eGFR, the influence on the cystatin C concentration from non-GFR determinants is of the same magnitude as the influence of non-GFR determinants on the creatinine concentration.

Among several proposed non-GFR determinants influencing the plasma cystatin C level, few have been established conclusively. Thyroid disease and
systemic high-dose treatment with corticosteroids have been demonstrated to directly influence the cystatin C concentration, whereas there are conflicting results for the influence of age, sex and BMI. Inflammatory states have been shown not to directly influence the cystatin C level (7, 33-38).

More than 15 eGFR formulae based on cystatin C levels have been published (12). The CKD-EPI collaboration published an equation to estimate GFR from cystatin C in 2008, which was re-expressed for use with standardized cystatin C in 2011 (39, 40). A new CKD-EPI cystatin C equation developed in 2012 (includes age and sex) and the CAPA equation from 2014 (includes age) are the main equations applied today (41, 42). Thus, cystatin C has the advantage of not requiring information on ethnicity for eGFR. Although cystatin C based eGFR is more accurate in certain specific circumstances including extremes of muscle mass or strict vegetarian diet, the overall accuracy of cystatin C based eGFR is not consistently higher compared with creatinine based eGFR (7, 42). In the KDIGO guidelines, measuring cystatin C is recommended in adults with eGFR from creatinine of 45–59 ml/min/1.73 m² BSA without markers of kidney damage if confirmation of CKD (GFR <60 ml/min/1.73 m² BSA) is required (5).

Creatinine and cystatin C combined
Combining cystatin C and creatinine to increase accuracy of eGFR was proposed in 1999 (43). When cystatin C and creatinine are used in conjunction to estimate GFR, P30 values are close to 90%, providing the most accurate estimation of GFR from endogenous filtration markers in clinical use (7, 44). This implies that cystatin C and creatinine have different non-GFR determinants.

Several different equations combining cystatin C and creatinine have been published, including the 2012 CKD-EPI creatinine-cystatin C equation (includes age, sex and ethnicity) currently recommended in the KDIGO guidelines (5, 42). It is noteworthy that when generating and validating the CKD-EPI creatinine-cystatin C equation, the authors observed a comparable accuracy in estimating mGFR when the arithmetic mean of eGFR from creatinine and cystatin C was evaluated.

Prevalence of CKD
The introduction of simple equations to estimate GFR from endogenous filtration markers, age, sex and ethnicity has enabled researchers to investigate the prevalence of CKD in large general population samples. Normal GFR in a healthy young adult Caucasian is about 125 ml/min/1.73 m² BSA
(7). GFR declines with normal ageing starting at ages 30-40 years and accelerating after the ages of 50-60 years (45, 46). The rate of decline has not been determined conclusively but available evidence suggests a rate of decline starting below 1 ml/min/year and then accelerating with increasing age (45, 46). Thus, the introduction of uniform thresholds of eGFR for diagnosis and stratification of CKD regardless of age with the 2002 KDIGO guidelines, may have led to an overestimation of CKD prevalence in the general population (16). Especially since CKD stage 3 does not require signs of kidney damage. The global prevalence of CKD, when the KDIGO classification is applied, is in the range 8-16% (1).

**Risk factors for CKD**

CKD and CVD display a bi-directional association, each being associated with a high magnitude risk to develop the other condition (47, 48). CKD and CVD share several conventional risk factors including obesity, diabetes, hypertension, hyperlipidaemia and smoking (1, 49-56). The pathophysiological mechanisms of atherosclerosis and glomerulosclerosis are to some extent parallel processes, both associated with inflammation (57-60). It is noteworthy that diabetes mellitus, hypertension, obesity and CVD are all part of the rising global burden of non-communicable disease (1, 54). Increasing age is associated with reduced GFR and increasing prevalence of CKD (61).

Family history of CKD is relevant not only for monogenic disorders like autosomal dominant polycystic kidney disease but also in reflecting genetic predisposition to develop CKD in general (62-64). Specific forms of CKD including diabetic nephropathy and the most common form of glomerulonephritis in developed countries (IgA nephropathy) are known to be influenced by hereditary factors (65, 66). Socioeconomic disadvantage is another factor which increases the risk of CKD (67-70).

In the case of CVD, an origin in childhood and adolescence conferring increased risk in later stages of life is well established (71), but early risk exposure is often associated with later accumulation of further risk (72). Markers of risk prior to adulthood for subsequent CKD are less well described, since few studies have evaluated risk factors for CKD from a life-course perspective. However, prematurity, low birth weight, rapid weight gain in childhood and obesity in childhood or adolescence have been associated with future CKD in adulthood (73-75). Paper I in this thesis evaluates predictors in adolescence of end-stage renal disease (ESRD) in middle-aged men.
Inflammation and CKD
After CKD onset, the prevalence of elevated inflammatory markers is high (76). Associations between different markers of inflammation and incident CKD have been demonstrated in several cohorts of middle-aged or older individuals, including the ARIC (Atherosclerosis Risk in Communities) study and the Multi-Ethnic Study of Atherosclerosis (MESA) (77-79). Although associations between markers of inflammation before adulthood and CVD in later life have been identified, similar associations with CKD have not been evaluated (80). This is addressed in paper I of this thesis which include evaluation of associations between erythrocyte sedimentation rate (ESR) in late adolescence and ESRD in middle-age.

Infections and CKD
Acute onset kidney disease following infectious events has been recognized since the 19th century (81). Certain acute infections are associated with increased risk of glomerulonephritis possibly leading to CKD. However, the incidence of post-infectious glomerulonephritis, including typical post-streptococcal glomerulonephritis, has declined during the recent decades and is now considered low in western countries (82, 83). From a global perspective, chronic infections including HIV, hepatitis B and hepatitis C are important risk factors for CKD (1).

Little is known about the possible long term risk of CKD following acute infectious episodes. In a registry based study from Taiwan, hospital admission with pneumococcal pneumonia was associated with an increased risk of ESRD during up to 13 years of follow-up (84). Increased risk of mortality, CVD and congestive heart failure are known to persist for more than 10 years following pneumonia (85-90). These adverse outcomes are associated with the magnitude and persistence of the inflammatory response to pneumonia (90-92). This suggests that inflammation and vascular disease, resulting from serious infections, may trigger not only acute post-infectious glomerulonephritis, but also possibly initiating processes resulting in delayed kidney disease diagnosed several years later (81). Paper III in this thesis evaluates possible long term increased risk of CKD following hospital admission with pneumonia.

Consequences of CKD
The increased risk of predominantly cardiovascular mortality and the high prevalence of CVD in ESRD have been recognized since the nineteen seventies (93). In the late eighties and early nineties, increased risk of mortality
associated with mild to moderate reduction of GFR was described in high-risk groups including elderly patients and those with hypertension, diabetes, myocardial infarction, congestive heart failure or stroke (94-99). Given the close relationship between CVD and CKD, it is not surprising that at this time, increased risk associated with elevated serum creatinine was supposed to be an indication of generalized atherosclerosis due to hypertension. From the late nineties and onwards, several studies supporting an association between elevated serum creatinine and increased morality in the general population were published (100-102). In 2004 limitations of previous studies including dichotomous measures of kidney function (serum creatinine or eGFR), few individuals with CKD or lack of information on comorbidity, were addressed by Go et al in a landmark paper establishing an independent, graded association between a reduced estimated GFR and the risk of death, cardiovascular events, and hospital admission in a large, community-based population (48). Thus, although CKD and CVD share risk factors and have similarities in their pathophysiological processes, CKD is an important independent risk factor for CVD which confers a two to four times increased risk after adjustment for conventional risk factors (48, 103, 104).

In addition to reduced GFR, the second hallmark of CKD is albuminuria and both are independently associated with increased mortality and CVD with a multiplicative effect when combined (105).

An individual who has developed CKD is not only at a substantially increased risk of mortality and CVD but may also face ESRD which has a major impact on quality of life and life expectancy (106). The cost for treatment of ESRD in developed countries constitutes 2-3% of healthcare expenditure although these patients account for only 0.1-0.2% of the total population, and less severe CKD is associated with even higher economic costs (1).

**eGFR and associations with adverse outcomes**

The introduction of simple formulae to estimate GFR facilitated a large scientific interest in associations between eGFR and adverse outcomes including all-cause mortality, cardiovascular mortality and CVD. This accelerated further after the publication of the study by Go et al. in 2004 (48, 107). At the same time numerous studies were conducted validating eGFR formulae in different populations, comparing the performance of different eGFR formulae using serum creatinine and/or cystatin C or introducing alternative formulae to estimate GFR. This is illustrated by the number of scientific
publications identified in MEDLINE when searching for papers with ‘estimated GFR’ or ‘MDRD’ or ‘CKD-EPI’ mentioned in title or abstract (figure 1).

![Figure 1](image_url)

*Figure 1. Publications in MEDLINE utilizing eGFR.*

It is important to remember that eGFR equations are not optimized for evaluation of associations between GFR or the endogenous filtration marker with adverse outcomes. This is illustrated by the fact that eGFR from cystatin C consistently has a higher magnitude association with mortality than the more precise eGFR based on creatinine and cystatin C combined and that decline in mGFR is not more predictive of mortality than decline in eGFR from cystatin C or eGFR from creatinine (42, 44, 108-112). Associations between eGFR and adverse outcomes including mortality and CVD can be confounded by potential associations between age, sex and ethnicity incorporated in the equations and the outcome of interest. Possible residual confounding from non-GFR determinants of the filtration marker not accounted for by the proxy measures included in the equation is an additional problem (113). Among these non-GFR determinants, low creatinine production, mainly due to reduced muscle mass (sarcopenia) associated with increased mortality risk, is of particular importance. However, for serum cystatin C associations with all-cause mortality independent of mGFR,
have not been firmly established. This is largely due to difficulties in interpreting results of previous studies given the high collinearity between cystatin C and mGFR in statistical models (114-116).

Paper II in this thesis evaluates associations between cystatin C and all-cause mortality independent of mGFR, avoiding influence from multicollinearity by comparing the performance of nested models.

**Sarcopenia**

Muscle mass and muscle strength have important implications for the understanding of GFR estimation from serum creatinine and associations between eGFR calculated from serum creatinine and adverse outcomes. Low muscle mass with reduced creatinine production reduces serum creatinine concentrations resulting in overestimation of GFR. Sarcopenia has also emerged as a powerful independent marker of increased mortality risk valid as well in the general population as in populations with CVD or CKD (117-120). This contributes to the J-shaped or U-shaped association observed between serum creatinine as well as eGFR from serum creatinine and all-cause mortality risk (121-124). Paper II in this thesis includes investigation of the functional form of associations between serum creatinine and all-cause mortality with and without adjustment for other measures of GFR.

Low muscle strength, predominantly measured as hand grip strength, has notable associations with increased all-cause mortality (125, 126). Low grip strength is also one of several clinical characteristics that identify the clinical syndrome of frailty which is characterized by increased vulnerability to endogenous and exogenous stressors (127). Muscle strength has a moderate correlation with muscle mass (128). However, the decline in muscle strength with increasing age is not fully explained by loss of muscle mass (128).

Paper IV in this thesis evaluates possible confounding and hypothesized effect modification by low grip strength on the association between creatinine-based eGFR and all-cause mortality.
**Aims**

The two overall aims of this thesis are to advance the understanding of how different exposures during the life-course can influence the risk of CKD in adult life and to deepen the understanding of how markers which are used to identify reduced GFR in CKD indicate risk of future adverse events.

The specific aims of the four papers of this thesis are:

Paper I. To examine markers of health and function in adolescence available from conscription and socioeconomic indicators, for associations with ESRD in middle-age.

Paper II. To assess whether cystatin C is associated with mortality independent of mGFR. A secondary aim is to evaluate the utility of combining cystatin C and creatinine to predict mortality risk.

Paper III. To investigate whether pneumonia in adult life requiring inpatient care results in a persistent raised risk of subsequent CKD.

Paper IV. To evaluate whether grip strength modifies the association between eGFR calculated from serum creatinine using the CKD-EPI equation and all-cause mortality.
Material and methods

Data sources

National population and health registers
Swedish national population and health registers hold a wealth of information which, after approval from an ethical review board, can be utilized for research purposes (129). The unique personal identification number has been issued to all residents of Sweden at birth or immigration since 1947, which enables linkage across registers and allows largely complete follow-up of each individual over time. This provides excellent conditions for register-based research which, together with the other Nordic countries, are unique in an international perspective (129).

Paper I, paper II and paper III of this thesis utilized data from national registers, and these registers are briefly described below.

The Total Population Register
Since 1968 the government agency Statistics Sweden maintains the Total Population Register which details age, sex, births, deaths, immigration, emigration, marital status, citizenship, country of birth and postal address. The source of the data is from several other registers, including the Population Register held by the Swedish National Tax Agency from which updates are received continuously (130).

The Swedish Population and Household Censuses
Population censuses have been conducted in Sweden since the mid eighteenth century and household censuses since the beginning of the twentieth century. Between 1960 and 1990 coordinated population and household censuses were held every five years.

The Cause of Death Register
Efforts to collect cause of death data on a population level have been made in Sweden beginning in 1751. Electronic records are available from 1952 and onwards in the Swedish Cause of Death Register held by the Swedish National Board of Health and Welfare. This is a virtually complete register of all deaths in Sweden with the underlying cause of death and contributing causes of death coded according to the International Statistical Classification of Diseases and Related Health Problems (ICD) from the World Health
Organization (WHO). In contrast with other Swedish national registers the Cause of Death Register uses the international version of this classification system. Coding in the register follows the different revisions of ICD with codes registered according to ICD-6 beginning in 1951, ICD-7 beginning in 1958, ICD-8 beginning in 1969, ICD-9 beginning in 1987 and from 1997 and onwards ICD-10 (131).

The Swedish Prescribed Drug Register
The Swedish Prescribed Drug Register covers all prescribed medications collected by patients since July 1, 2005 (132).

The National Patient Register
The National Patient Register (NPR) is held by the Swedish National Board of Health and Welfare. The register details inpatient diagnoses and procedures beginning in 1964, with complete national coverage since 1987. Since 2001 the register also covers outpatient care including day surgery and psychiatric care from both private and public caregivers. Primary care is not covered in the NPR. Reporting to the register is mandatory for all health care providers (133, 134).

Diagnoses are coded according to the Swedish version of ICD-7 beginning in 1964, ICD-8 beginning in 1968, ICD-9 beginning in 1987 and finally ICD-10 from 1997 and onwards (with the exception of the county of Skåne where the introduction of ICD-10 was delayed by one year). The quality of the register has been reviewed by Ludvigsson et al. who also provide a detailed description of the register including the coding of medical procedures (134).

The Swedish Military Conscription Register
Beginning in 1901 and continuing throughout the twentieth century, assessment for compulsory military service was mandatory by law for all male Swedish citizens. The Swedish Military Conscription Register was established in 1952 and includes information collected during the conscription assessment. In the seventies less than 4% of men with Swedish citizenship were excluded from the assessment due to either a severe chronic medical condition or handicap documented in a medical certificate or incarceration. The initial conscription assessment was at that time conducted during two days and included evaluation by a physician and the collection of measures for medical evaluation, evaluation of physical performance, an assessment
of cognitive function and psychological profile including evaluation by a psychologist (135).

The Conscript Cohort
The study population in paper I and paper III is a cohort of all male residents in Sweden born from 1952 to 1956 with records in the Swedish Military Conscription Register (n=284,198). These men, representing 97-98% of that Swedish male birth cohort were followed until 31st December 2009 using Swedish national registers with linkage through the Swedish unique personal identification number. The conscription assessments were between 1970 and 1975 and the majority at ages 18 and 19 years, with a small number after this time at later ages.

Men with inconsistencies in their data such as incorrect personal identification number or uncertain vital status, were excluded (n=2,564). A further 225 men were excluded from the analysis due to improbable measures at the conscription assessment: height less than 144 cm (n=39); BMI below 15 (n=134); weight above 178 kg (n=9); systolic blood pressure below 50 or above 230 mm Hg (n=33); and diastolic blood pressure below 30 or above 135 mm Hg (n=12). Men who did not complete the mandatory conscription examination due to chronic illness, disability or lack of Swedish citizenship (n=16,458), were also excluded. This resulted in a cohort of 264,951 men.

The Total Population Register provided dates of birth, death and emigration. The Swedish Military Conscription Register provided information on the baseline conscription examination including BMI, blood pressure, ESR, erythrocyte volume fraction (EVF), dip-stick proteinuria, physical working capacity (assessed using an electronically braked ergometer) and a cognitive function score. Head of household’s occupation and household crowding measured as person-per-room ratio, when participants were children, was accessed through the 1960 Population and Housing Census. Routinely collected health data in the NPR provided diagnoses defined using both procedure codes, and ICD-8, ICD-9 and ICD-10 disease codes. Diabetes and hyperlipidaemia diagnoses were additionally detected through the Swedish Prescribed Drug Register.

The mGFR Cohort
Paper II in this thesis utilizes a consecutive patient series of 1,286 Swedish residents aged 18 years or older referred to the Department of Laboratory
Medicine at Örebro University Hospital between 2004 and 2010 for measurement of GFR by plasma iohexol clearance, with sufficient serum to determine cystatin C and creatinine (>99%). Major indications for referral included a CKD diagnosis and follow-up, evaluation for treatment with drugs cleared by the kidneys (including chemotherapeutic drugs), evaluation of potential kidney donors and follow-up of patients treated with lithium. At this time there was no international standardization for measurements of cystatin C and thus the local application of eGFR equations developed elsewhere was problematic. The patient series was originally assembled at the laboratory for quality assurance purposes and in order to be able to develop a local equation to calculate eGFR from cystatin C. After approval from the Ethical Review Board of Uppsala Sweden, these patients formed the basis of the study cohort in paper II. The individuals in the cohort were characterised using the NPR and followed using the Swedish Cause of Death Register until 31st December 2012.

The United Kingdom Household Longitudinal Survey

Paper IV in this thesis utilizes data from the UK Household Longitudinal Survey (UKHLS). This large nationally representative longitudinal panel survey is following members in all ages from about 40,000 households in the UK since 2009-2010. The survey covers a broad range of themes such as family life, education, employment, finance, behaviour, health and well-being. Annual interviews during a visit by an interviewer or on-line interviews are performed in overlapping 2-year waves (8). Although designed to be broadly representative for the general UK population, the UKHLS only samples private households and thus does not include people living in institutions, including care homes (136).

The British Household Panel Survey (BHPS) established in 1999-2001 has followed the members of about 9,000 UK households up to 2008 when the remaining study participants were invited to join the UKHLS at wave 2 (2010-2012) (8).

The study population in paper IV is a subsample of the UKHLS general population sample which were visited by a trained nurse to collect anthropometric and health measures including non-fasting blood samples. This was performed in average 5 months after the wave 2 interview for non-BHPS participants and the wave 3 interview for the BHPS sample (8, 137).

Data from the UKHLS are available through the UK Data Archive for researchers who register and provide details of their research project to demonstrate that their work is in the public interest. This type of access was
used to obtain the data for paper IV. More sensitive data such as day and month of birth, detailed country of birth, more detailed information on geographic location, etc. require a special licence or secure access granted under specific conditions.

**Statistical methods and concepts**

**Cox regression**

Cox proportional hazards regression for survival-time (time-to-event) outcomes, introduced in 1972 by Sir David Cox, is one of the most popular regression techniques for analysis of survival data with censored failure times (138, 139). The logarithm of the incidence rate (hazard rate) is modelled as a multiple linear regression on a set of explanatory variables, with the baseline incidence rate being an ‘intercept’ term that varies with time. A major advantage with the Cox model is its semi-parametric construct where the nonparametric baseline hazard does not need to be estimated to calculate the coefficient for the effect of an explanatory variable (parametric component). However, this requires an important **assumption of proportional hazards**. At any given time, the ratio between the hazard of exposed to unexposed (or the ratio between the hazards before and after an increase of one unit in a continuous exposure) must be constant. In other words, the proportional hazards assumption translates into a constant effect of a given exposure throughout the follow-up time (140).

Another important assumption for the Cox model is the **assumption of non-informative censoring**, which states that the mechanisms giving rise to censoring must not be related to the probability of an event occurring. A third assumption is the **assumption of a linear relation between each independent variable and the log hazard**.

**Time dependent exposures and time dependent effects**

In paper III of this thesis data were analysed using Cox regression. The exposure pneumonia was not present at start of follow-up. During follow-up a proportion of the study population were hospitalized with pneumonia and from that time considered exposed. The proportional hazards assumption was found to be violated for the exposure pneumonia, implying an effect on survival time without CKD diagnosis that was not constant during follow-up after pneumonia (141).

This was resolved by splitting follow-up time at time of pneumonia and at specified time points after pneumonia. Thus, the model estimated the
hazard ratio (HR) for CKD after exposure to pneumonia in different time-periods.

In paper II the proportional hazards assumption was violated for the co-variates age and pre-existing cancer in the Cox regression model. This was addressed by an internal stratification by cancer status and tenths of the age distribution, which allows for different baseline functions in these different strata.

**Logistic regression**

Binary logistic regression developed by Sir David Cox in 1958, models the logarithm of the odds (logit) of a binary outcome as a linear function of the explanatory variables. The regression coefficient for each explanatory variable represent the influence of a unit increase in that variable on the logit of the binary outcome, holding all other explanatory variables constant. Exponentiating the coefficient returns the odds ratio (OR) (142, 143).

Besides the obvious premise of a binary outcome, there are four main assumptions that need to be met for the correct application of binary logistic regression. *Assumption of observational independence.* Thus, the observations should not come from repeated measurements or matched data. *Assumption of linearity of independent variables and the log odds.* Continuous independent variables should have a linear relationship with the log-transformed outcome. *Assumption of the absence of multicollinearity among independent variables,* and finally, *assumption of absence of strongly influential outliers* (143, 144).

**Conditional logistic regression**

Conditional logistic regression is an extension of logistic regression suitable for matched data. In a case-control setting the method models the log of the odds (logit) of being a case as a linear function of the explanatory variables and a constant term for each stratum (risk-set). Exponentiating the regression coefficient for a variable translates into an odds ratio as in unconditional logistic regression described above. However, conditional logistic regression requires an additional *assumption of equal odds ratios for each explanatory variable in all strata* (145).

In a case-control setting, the number of cases (positive outcomes) in each stratum (risk-set) is predetermined by the study design. Conditional logistic regression fits a logistic model that explains why a specific individual had a positive outcome within each stratum, conditional on that only one of the individuals in each stratum have a positive outcome. Since the comparison
is within each stratum, conditioned on the number of positive outcomes (cases), the strata specific intercepts (constants) in the model cancel out and remain unestimated (146).

Multivariable fractional polynomial method
The regression models applied in this thesis all assume a linear relationship between continuous independent variables and the measure being modelled (the log hazard or the log odds). However, the shape of the association (functional form) may be non-linear. This can be addressed by categorizing a continuous variable or by transforming the continuous variable with the aim to achieve this linear association.

In paper I and paper III, continuous measures were all categorized apart from single measures used for adjustment only, where a linear relationship was assumed. In paper II and paper IV, multivariable fractional polynomials were used to account for possible non-linearity in the association for all continuous measures.

The multivariable fractional polynomial method (MFP) was applied as described by Royston (147, 148). Transformations of continuous variables were selected from a fixed set including exponentiating the variable by the power of -2, -1, -0.5, 0, 0.5, 1, 2, 3 were 1 indicate no transformation and 0 is equal to logarithmic transformation. Both first-degree (one-term) fractional polynomial (FP1) functions and second-degree (two-term) fractional polynomial (FP2) functions combining two transformations of the variable, could be selected. The MFP method is a complex iterative process that will be only broadly outlined here.

In the first cycle, the FP2 transformation of each variable producing the best model-fit is chosen. If this transformation produces a statistically significantly better model-fit as compared to including the untransformed variable in the model, the transformation is retained. However, this FP2 transformation is then compared with the best-fitting FP1 transformation, and only if the FP2 transformation produce a statistically significantly better model fit, the FP2 transformation is selected, otherwise the FP1 transformation is adopted. In the second cycle, the covariates are included in the model with the transformations chosen in the first cycle. All covariates are then examined again, in descending order of statistical significance, selecting the transformation as described in the first cycle, but now in a model containing the transformed variables as choses in the first cycle. This is repeated until the same transformations are selected in two consecutive steps (147, 149). The MFP method will also select which covariates to retain in
the model. However, we have not applied any data driven selection of variables in our models, but retained all pre-specified variables.

Optimizing a model to fit the data may result in overfitting where random ‘noise’ in the specific dataset used is included in the model. This results in poor predictions when the model is applied to other data. Considering the risk of overfitting, we used a 20% significance level for transformation of covariates included for adjustment and a more conservative level of significance for transformation of the predictor of interest (10% in paper II and 5% in paper IV) (147).

**Incidence density sampling**

In paper III we applied incidence density sampling without replacement controls. The name refers to controls being selected at each time-point an incident event (case) occurs. In a case control setting, controls are selected for each case, that are still at risk of becoming a case at the time the case occurred. In paper III, we conducted a sensitivity analysis restricting the study population of men from the Conscription Cohort to only men who had at least one hospital admission during follow-up. Then a matched nested cohort was created by matching each man with pneumonia (still at risk of CKD) with five unexposed men (at risk of pneumonia and at risk of CKD) by birth year, month and year of discharge. In incidence density sampling a control can later become a case. Thus, controls were censored at the time of first pneumonia diagnosis (150). This provided a model where future risk of CKD among men without diagnosed CKD, admitted to hospital with pneumonia for the first time, could be compared with men of the same age, without diagnosed CKD, admitted to hospital during the same calendar month, but without a prior or current episode of pneumonia.

**Net reclassification improvement**

Net Reclassification Improvement (NRI) is a measure of model discrimination typically applied to quantify the added contribution of a new marker added to a risk prediction model. It was developed to address the shortcomings of standard methods to evaluate whether a new marker of risk improves present models of risk prediction. A new marker can have a statistically significant association with the outcome when added to a model without improving risk prediction to a meaningful extent, especially in large samples. The C-statistic and AUC are instead too conservative with only very small changes in their magnitude when a new predictor is added once the model contains a few good predictors (151).
The NRI is the sum of two components, the event-NRI (the net proportion of patients experiencing an event reclassified to a higher risk) and the non-event-NRI (the net proportion of patients not experiencing an event reclassified to a lower risk). When established categories of risk for the outcome of interest exist, these categories can be used in the evaluation of reclassification. However, a category free NRI is considered more objective and allows comparison of results between studies. The NRI should be presented with its two components and with 95% confidence intervals (CI). The 95% CI are attained by bootstrapping (151, 152).

We applied category free NRI in paper II to evaluate potential improvement in model performance when adding creatinine, cystatin C or mGFR to models where one or two of the other markers were already included.

**Multicollinearity**

Multicollinearity arises when the independent variables in a model are moderately or highly correlated. If two independent variables are highly correlated then it can be difficult to assess the specific association of one with the dependent variable, whereas inference for multivariable analysis is based on the assumption that all independent variables are uncorrelated (153-155).

Multicollinearity leads to biased coefficient estimation and a loss of power. Collinearity increases the standard errors of regression coefficients, causing wider CI, which promote rejection of significant test statistics. Multicollinearity in a sense leads to mixing of effects from different explanatory variables and in turn imprecise estimates of regression coefficients, possibly influencing the direction and magnitude of associations (153). However, multicollinearity does not limit evaluation of improvement in model performance for a predictor when comparing nested models (154). This approach is utilized in paper II of this thesis to evaluate statistically significant independent associations with all-cause mortality for cystatin C and creatinine when adjusting for mGFR.

Multicollinearity can be detected by variance inflation factors (VIF). The VIF for each independent variable in a regression model is the factor by which the variance of its regression coefficient is inflated. The VIF is calculated from the R-squared value (percentage of the variation explained) obtained when regressing the variable on the other predictors in the model. Lower thresholds indicating concern can be chosen, but a VIF above 10 is considered a strong indication of problematic multicollinearity effects (153-155). The reciprocal of VIF, tolerance, is interpreted as the proportion of
the variance in an independent variable not explained by the other independent variables (154, 155).

Effect modification and interactions
The effect of an exposure on an outcome is not always homogenous across a population. Different subgroups defined by the level of another variable may have different associations between the exposure and the outcome. This heterogeneity of effect across subgroups can be described as effect modification. The characteristic described by the variable identifying the subgroups is said to modify the effect of the exposure on the outcome (156).

It is important to remember that there are numerous possible subgroups which might be created according to characteristics of the study participants and each characteristic may also be used in numerous different ways to create subgroups. Exploratory examination of heterogeneity in the effect of an exposure across all of these potential subgroups is prone to produce spurious findings of effect modification. Thus, when investigating possible effect modification this should be biologically plausible and hypothesis based when designing the study (156).

In regression models including in the papers of this thesis, multiplicative interactions are assessed where there is evidence of different magnitude associations in different strata and the association of the interaction term remains statistically significant after adjustment for the main effects.

In paper IV the aim was to describe the association between eGFR and all-cause mortality in individuals with different levels of grip strength. Thus, the interaction was described as effect modification which was illustrated by stratifying the model in three strata of grip strength, and therefore not evaluating the association between grip strength and the outcome. An independent variable cannot be both a confounding factor and an effect modifier in the same model.
Individual study designs and methods

Paper I – Predictors in adolescence of ESRD in middle-aged men

Study design
We conducted a register-based nested case-control study using the Conscription Cohort to examine markers in adolescence for their associations with ESRD in subsequent adult life, up to 40 years later. The main markers were hypertension, ESR as a marker of inflammation, including from infections, dip-stick proteinuria as a signal of early kidney damage, and BMI. These are readily available markers of health (although another marker of inflammation would be used in a more modern setting) and development in adolescence, which theoretically could be associated with biological processes relevant to risk of future ESRD. To avoid results being driven by undiagnosed kidney disease in adolescence, we began follow-up in 1985, which is at least 10 years after the conscription examinations.

Table 2. Diagnostic codes at conscription examination. ICD-8 Swedish version

<table>
<thead>
<tr>
<th>Disease group</th>
<th>ICD-8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urogenital disease</td>
<td>400.30, 403.99, 404.99, 580-607, Excluding 593.22</td>
</tr>
<tr>
<td>Kidney and urinary tract malformations</td>
<td>753</td>
</tr>
<tr>
<td>Diabetes</td>
<td>250</td>
</tr>
<tr>
<td>Other endocrine, metabolic or nutritional disorders</td>
<td>240-279, Excluding 250, 277</td>
</tr>
<tr>
<td>Systemic inflammatory disease</td>
<td>446, 712, 734</td>
</tr>
<tr>
<td>Cardiovascular disease and malformations</td>
<td>390-458, 746-747, Excluding 400.00, 400.30, 401.99, 403.99, 404.99, 446, 458.00</td>
</tr>
</tbody>
</table>
diagnosis of stage 5 CKD from 1985 to 2009 was identified in the NPR using the codes displayed in table 3.

### Table 3. Definition of ESRD. ICD-8/9/10 Swedish version. Surgical and medical procedure codes using Swedish classifications

<table>
<thead>
<tr>
<th>ESRD</th>
<th>ICD-8</th>
<th>ICD-9</th>
<th>ICD-10</th>
<th>Procedure codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>CKD stage 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kidney transplantation</td>
<td>V42A</td>
<td>Z94.0</td>
<td>6070</td>
<td>KAS10, KAS20</td>
</tr>
<tr>
<td>Dialysis general</td>
<td>Y29.01</td>
<td>V56</td>
<td>Z99.2</td>
<td>6070, KAS10, KAS20, 6074, PBL10, PBL20, PBL30, PBL99, PEL10, PFL90, DR014, DR016,</td>
</tr>
<tr>
<td>Haemodialysis</td>
<td>V56A</td>
<td>Z49.1</td>
<td></td>
<td>9200, V9200, 9212, V9212, V9223, V9223, V9507, 8872, 8873, 8874, PBL10, PBL20,</td>
</tr>
<tr>
<td>Peritoneal dialysis</td>
<td>V56W</td>
<td>Z49.2</td>
<td></td>
<td>9214, V9531, TJA33, JAK10, DJ008, DR012, DR013, DR024, QF006</td>
</tr>
</tbody>
</table>

**Statistical analysis**

Controls were selected within the cohort with a ratio of 10 per case of incident ESRD and matched for birth year, vital status and county. Men with a cardiovascular diagnosis (table 2) at the conscription assessment were excluded at this stage. Differences between characteristics of cases and controls were analysed using the t-test for continuous variables and Pearson's chi-squared test for categorical variables.

Conditional logistic regression was used to examine the association with ESRD for proteinuria, hypertension, ESR and BMI. ESR was adjusted for EVF in all analyses. The adjusted model included all of the above measures, with adjustment for physical working capacity, cognitive function score; as well as head of household’s occupation and household crowding from the 1960 census.
Paper II – Measured glomerular filtration rate does not improve prediction of mortality by cystatin C and creatinine

Study design
We conducted a cohort study using the mGFR Cohort to assess whether cystatin C is associated with mortality independent of mGFR and to evaluate the utility of combining cystatin C and creatinine to predict mortality risk. Pre-existing diagnoses of diabetes, cancer, CVD and ESRD were identified in the NPR using ICD-codes and procedure codes listed in table 4. Patients with an earlier diagnosis of ESRD (n=129) were excluded, resulting in a study population of 1,157 individuals.

Table 4. Definition of disease groups and End Stage Renal Disease. ICD-8/9/10 Swedish version. Surgical and medical procedure codes in Swedish classifications

<table>
<thead>
<tr>
<th>Disease</th>
<th>ICD-8</th>
<th>ICD-9</th>
<th>ICD-10</th>
<th>Procedure codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diabetes</td>
<td>250,00-250,99</td>
<td>250A-250X</td>
<td>E10.0-E14.9</td>
<td></td>
</tr>
<tr>
<td>Cancer</td>
<td>140,01-209,99</td>
<td>140A-208X</td>
<td>C00.0-C97.9</td>
<td></td>
</tr>
<tr>
<td>CVD</td>
<td>390,00-458,99</td>
<td>391A-391X</td>
<td>I01.0-I01.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>392A</td>
<td>I02.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>393A-453X</td>
<td>I05.0-I82.9</td>
<td></td>
</tr>
<tr>
<td>ESRD</td>
<td>Y29,01</td>
<td>V56</td>
<td>Z94.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>V56A V56W</td>
<td>Z92.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>V45B</td>
<td>Z99.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Z49.0-Z49.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6070, KAS10, KAS20, 9200, V9200, 9212, V9212, 9223, V9223, V9507, 8872, 8873, 8874, PBL10, PBL20, PBL30, PBL99, PEL10, PFL90, DR014, DR016, DR017, DK001, SP312, SP322, SP332, SP342, SP352, SP392, 9214, V9531, TJA33, JAK10, DJ008, DR012, DR013, DR024, QF006</td>
<td></td>
</tr>
</tbody>
</table>

Statistical analysis
Filtration markers and mGFR are highly collinear and multicollinearity may influence the magnitude and the statistical significance of each measure’s association with mortality when included in the same models. However, multicollinearity effects will not limit the evaluation of the independent information on mortality risk added by each marker when nested models are compared (155).

We examined associations of baseline markers of kidney function with mortality using Cox regression. Time at risk was calculated from the date
of the iohexol clearance examination to the date of death, emigration or end of follow-up, whichever occurred first. All adjusted models included age, sex, body mass index (BMI) and pre-existing diagnoses of CVD, diabetes and cancer. The final adjusted Cox regression models further allowed for different baseline functions by cancer status and tenths of the age distribution. HR were calculated comparing the median value of quintiles of cystatin C\(^1\), creatinine\(^1\) and mGFR with the median of the approximate range of normal values for each marker (age<50 years and averaging men and women): creatinine 75 mmol/L, cystatin C 0.925 mg/L and mGFR 100 ml/min/1.73m\(^2\) BSA.

Linearity of the relationship between the log hazard function and continuous variables was assessed by applying the multivariable fractional polynomial (MFP) method (147, 148). The variables were pre-specified and no data-dependent selection was performed. No exclusion of statistically suggested extreme outliers was performed, as they were considered biologically plausible. The functional form of BMI, the only continuous potential confounder in the model, was determined by fitting an MFP model with a 20% level of significance for variable transformation and for the markers of kidney function we used a 10% level of significance for variable transformation. When a 5% level of significance for transformation was employed, the same transformations were chosen for the markers of kidney function, suggesting that they were not forced by a complex data structure. BMI was modelled using a FP2 function with powers (-0.5, -0.5). mGFR was modelled as non-linear in the model including mGFR, creatinine\(^{-1}\) and the base model using a FP1 function with power (0). Creatinine\(^{-1}\) was modelled as non-linear using a FP2 function with powers (1, 2) when unadjusted; a FP2 function with powers (1, 1) when adjusted for the base model and a FP1 function with power (0) when adjusted for the base model and cystatin C\(^{-1}\) and when adjusted for the base model, cystatin C\(^{-1}\) and mGFR. Cystatin C\(^{-1}\) was modelled as non-linear using a FP1 function with power (-1) when adjusted for the base model; a FP1 function with power (0) when adjusted for the base model and mGFR and a FP1 function with power (-0.5) when adjusted for the base model and creatinine\(^{-1}\) and when adjusted for the base model, creatinine\(^{-1}\) and mGFR.

Sensitivity analyses included checking the adequacy of fractional polynomial functions for continuous variables in the final adjusted model using spline-based tests, whereby Chi-square tests and P values for joint testing of the additional restricted cubic spline terms were obtained. Knot positions were predefined and set at 5, 50, 95 centiles for the spline terms. The tests
suggested that splines would not improve the fit of the non-linear functions already selected by MFP. Further sensitivity analyses included exclusion of statistically suggested extreme outliers. The assumption of proportional hazards was assessed using a Grambsch–Therneau test of the scaled Schoenfeld residuals from a Cox model (157).

Pearson and Spearman correlations between markers of GFR are given in figure 2. Variance inflation factors (VIFs) and tolerance in a model including the markers of GFR and the other variables, are displayed in table 5.

**Table 5. Multicollinearity diagnostics for measures of GFR applied in the study and covariates**

<table>
<thead>
<tr>
<th>Marker</th>
<th>mGFR included VIF</th>
<th>mGFR included Tolerance</th>
<th>mGFR not included VIF</th>
<th>mGFR not included Tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>mGFR</td>
<td>9.62</td>
<td>0.10</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cystatin C⁻¹</td>
<td>6.15</td>
<td>0.16</td>
<td>3.46</td>
<td>0.29</td>
</tr>
<tr>
<td>Creatinine⁻¹</td>
<td>4.22</td>
<td>0.24</td>
<td>2.85</td>
<td>0.35</td>
</tr>
<tr>
<td>Age</td>
<td>1.72</td>
<td>0.58</td>
<td>1.48</td>
<td>0.68</td>
</tr>
<tr>
<td>Sex</td>
<td>1.31</td>
<td>0.76</td>
<td>1.14</td>
<td>0.87</td>
</tr>
<tr>
<td>BMI</td>
<td>1.14</td>
<td>0.88</td>
<td>1.07</td>
<td>0.94</td>
</tr>
<tr>
<td>CVD</td>
<td>1.39</td>
<td>0.72</td>
<td>1.39</td>
<td>0.72</td>
</tr>
<tr>
<td>Diabetes</td>
<td>1.23</td>
<td>0.82</td>
<td>1.21</td>
<td>0.83</td>
</tr>
<tr>
<td>Cancer</td>
<td>1.06</td>
<td>0.95</td>
<td>1.05</td>
<td>0.95</td>
</tr>
</tbody>
</table>

To assess the independent contribution of each marker of GFR to the performance of Cox regression models for all-cause mortality, we compared the performance of nested models where each kidney function marker was added to differently adjusted models (the base model included age, sex, BMI, pre-existing diagnoses of diabetes, CVD or cancer; base model and mGFR; base model and creatinine⁻¹; base model, creatinine⁻¹ and cystatin C⁻¹). Statistically significant differences in model fit were evaluated by applying the likelihood ratio test. The proportion of variance explained was evaluated by an adaption of the adjusted R² proposed by Royston for censored survival data (158). Discrimination was measured using Harrell’s concordance index (C-index) modified as proposed by Gönen and Heller (159). The category free NRI was calculated at 5 years (approximate mean follow-up time), and 95% CI for the NRI components were attained by bootstrapping with 1000 samplings (151, 152).
Figure 2. Correlation of filtration markers (inverted). Solid line is the linear regression fit and dashed line is the LOWESS smoother fit. Serum cystatin C is expressed in mg/L; serum creatinine in µmol/L and measured GFR in ml/min/1.73 m² BSA.
Paper III – Hospital admission with pneumonia and subsequent persistent risk of chronic kidney disease

Study design
We conducted a register-based cohort study using the Conscription Cohort to assess associations between hospital admission with pneumonia in adulthood and raised risk of subsequent CKD. ICD-8/9/10 codes, procedure

<table>
<thead>
<tr>
<th>Disease groups</th>
<th>ICD-8</th>
<th>ICD-9</th>
<th>ICD-10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pneumonia</td>
<td>480-486</td>
<td>480-486</td>
<td>J12-J18</td>
</tr>
<tr>
<td>Chronic kidney disease</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kidney failure</td>
<td>792.99</td>
<td>584-586</td>
<td>N17-N19</td>
</tr>
<tr>
<td>Acute kidney failure</td>
<td>580, 581-583</td>
<td>582-583</td>
<td>N00-N06</td>
</tr>
<tr>
<td>Glomerulonephritis</td>
<td>581</td>
<td>581</td>
<td>N04</td>
</tr>
<tr>
<td>Diabetic nephropathy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nephrosclerosis</td>
<td>584</td>
<td>587</td>
<td>I12-I13</td>
</tr>
<tr>
<td>Hypertensive kidney disease</td>
<td>400.30, 403.99, 404.99</td>
<td>403-404</td>
<td>I12-I13</td>
</tr>
<tr>
<td>Polycystic kidney disease</td>
<td>753.10</td>
<td>753B</td>
<td>Q61.1, Q61.2, Q61.3, Q61.9</td>
</tr>
<tr>
<td>ESRD</td>
<td></td>
<td></td>
<td>N18.5</td>
</tr>
<tr>
<td>Kidney transplantation</td>
<td></td>
<td>V42A</td>
<td>Z94.0</td>
</tr>
<tr>
<td>Dialysis in general</td>
<td>Y29.01</td>
<td>V56, V45B</td>
<td>Z99.2</td>
</tr>
<tr>
<td>Preparations for dialysis</td>
<td></td>
<td></td>
<td>Z49.0</td>
</tr>
<tr>
<td>Haemodialysis</td>
<td></td>
<td></td>
<td>Z49.1</td>
</tr>
<tr>
<td>Peritoneal dialysis</td>
<td></td>
<td></td>
<td>Z49.2</td>
</tr>
<tr>
<td>Sepsis</td>
<td>036.10, 038</td>
<td>036C, 038</td>
<td>A392, A40-A41, A499, R651, R57</td>
</tr>
<tr>
<td>Diabetes</td>
<td>250</td>
<td>250</td>
<td>E10-E11, E13-E14</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>390-458</td>
<td>391, 392A, 393-453</td>
<td>I01, I02, I05-I82</td>
</tr>
<tr>
<td>Hyperlipidaemia</td>
<td>272,00-272,09</td>
<td>272A-272E, 272X</td>
<td>E78.0-E78.5, E78.9</td>
</tr>
<tr>
<td>Chronic obstructive pulmonary disease</td>
<td>491-492</td>
<td>491-492</td>
<td>J41-J44</td>
</tr>
</tbody>
</table>

Exclusions at conscription assessment
codes and ATC-codes for prescribed medications from the Swedish Military Conscription Register, the NPR and in the Swedish Prescribed Drug Register are presented in table 6a and 6b.

<table>
<thead>
<tr>
<th>Disease groups</th>
<th>Procedure codes and ATC codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronic kidney disease</td>
<td></td>
</tr>
<tr>
<td>ESRD</td>
<td></td>
</tr>
<tr>
<td>Kidney transplantation</td>
<td>6070, KAS10, KAS20</td>
</tr>
<tr>
<td>Haemodialysis</td>
<td>9200, 9212, 9223, V9200/9212/9223, DR014/016/017</td>
</tr>
<tr>
<td>Creating arteriovenous fistula</td>
<td>8872-8874, PBL10/20/30/99, PEL10, PFL90</td>
</tr>
<tr>
<td>Central dialysis catheter, long term</td>
<td>SP312/322/332/342/352/392, V9507, DK001</td>
</tr>
<tr>
<td>Peritoneal dialysis</td>
<td>DR012/013/024, QF006, 9214, V9531</td>
</tr>
<tr>
<td>Insertion of peritoneal dialysis catheter</td>
<td>TJA33, JAK10, DJ008</td>
</tr>
<tr>
<td>Diabetes</td>
<td>A10 – Drugs used in diabetes*</td>
</tr>
<tr>
<td>Hyperlipidaemia</td>
<td>C10 – Lipid modifying agents*</td>
</tr>
<tr>
<td>Kidney biopsy</td>
<td>6082, KAB00, TKA00, AK062</td>
</tr>
</tbody>
</table>

* Indicate ATC code.

Men with diagnoses at conscription assessment indicating possible kidney dysfunction, including trace or positive dip-stick proteinuria; chronic inflammatory disease; men with CVD or malformations of the urogenital or cardiovascular system at conscription assessment were excluded. Further exclusions were made for conscription assessment outside ages 16–20 years. Finally, all men diagnosed with polycystic kidney disease at any time during follow-up were excluded. A total of 227,364 men (80% of the birth cohort) remained in the analytical sample.

After the initial conscription assessment, exposure to pneumonia in men without contemporaneous or previously recorded kidney disease was identified using hospital discharge records. Diagnoses of pneumonia were considered as separate episodes if the discharge diagnoses were more than 30 days apart.

The outcome CKD was defined as the first occurrence of the following conditions in the NPR; kidney failure, ESRD, glomerulonephritis, nephrotic syndrome, diabetic nephropathy, nephrosclerosis or hypertensive kidney disease. This wide definition was chosen since the majority of patients received their CKD diagnosis without a previous kidney biopsy (91.5%). ESRD was defined as long-term dialysis therapy, kidney transplantation,
surgical procedures creating long-term access for dialysis, or a diagnosis indicating stage 5 CKD.

From the baseline at conscription examination, established risk factors for CKD and measures previously associated with future ESRD in the same cohort, including socioeconomic conditions for the family of origin in the 1960 census (paper I) were included as covariates.

Statistical analysis
Characteristics of exposed and unexposed men were compared using the two sample T-test for continuous measures and Pearson’s chi-squared test for categorical measures (160, 161). Associations between hospital admission with pneumonia and subsequent CKD were evaluated using Cox regression with attained age as the underlying timescale (139). To reduce surveillance bias and to avoid detecting acute pneumonia-related kidney complications, the main follow-up began one year after the pneumonia episode. The study population was followed until date of first incident diagnosis of kidney disease, death, emigration or end of follow-up, whichever occurred first. Due to evidence of non-proportional hazards, follow-up time was divided into categories (≤5, >5-≤15, and >15 years), and for these periods there was no evidence of violation of the proportional hazards assumption, tested using scaled Schoenfeld residuals (162). The added variable version of the Gronnesby and Borgan test was used to test model fit for each separate period (163).

The use of routinely collected inpatient data may introduce selection and surveillance bias. A number of sensitivity analyses related to the main analysis was performed to further evaluate whether selection bias, surveillance bias, confounding from comorbidity or more severe pneumonia episodes associated with sepsis or acute kidney failure were likely to explain our findings. The analysis was repeated only considering pneumonia as the primary diagnosis at hospital discharge, adjusting for repeated pneumonia episodes, excluding acute kidney failure or diabetic nephropathy from the outcome and excluding men ever diagnosed with sepsis or men ever diagnosed with chronic obstructive pulmonary disease (COPD) from the study base. Additional adjustment was also undertaken for CVD, diabetes and hyperlipidemia, modeled as time-dependent covariates. We also restricted the outcome to ESRD, and collapsed the final two periods of follow-up time after pneumonia (>5-≤15 and >15 years) due to low numbers, in a model with additional adjustment for CVD, diabetes and hyperlipidaemia, modelled as time-dependent covariates.
Finally, to further address possible residual confounding by comorbidity and to reduce surveillance bias, the sample was restricted to include only men who had at least one hospital discharge diagnosis during adulthood. Incidence density sampling without replacement of unexposed was applied to create nested matched cohorts (150). Each man with pneumonia (at risk of CKD) was matched with five unexposed men (at risk of CKD and pneumonia) by birth year, month and year of discharge (index date), censored at the time of first pneumonia diagnosis (164). Thus, a man selected as control may later be included as a case. In these nested cohorts, additional adjustment was included for adult diabetes and CVD before or at the index date.

Paper IV – Grip strength modifies the association between estimated glomerular filtration rate and all-cause mortality

Study design
We conducted a cohort study using a subsample of the UKHLS general population sample to evaluate whether grip strength modifies the association between eGFR calculated from serum creatinine using the CKD-EPI equation and all-cause mortality (23). These individuals were visited by a trained nurse to collect anthropometric and health measures including non-fasting blood samples. Non-pregnant respondents aged 16 years or older, living in England, Scotland or Wales, who completed their interview in English were eligible for nurse visits, and among 35,937 eligible respondents 20,700 participated. Some 1,579 individuals who volunteered at the nurse visit that they were HIV positive, had hepatitis B or C, had a bleeding or clotting disorder, took anti-clotting medication excluding aspirin or had ever had a seizure, were not eligible to give blood. Another 4,688 respondents refused to give blood. Blood samples were successfully obtained and at least one blood-based biomarker determined in 13,107 individuals. After first excluding participants with missing data on any of the measures included in our analysis (n=1,476) and finally excluding participants with eGFR values outside the range of 15-120 ml/min/1.73 m² BSA (n=731), a total of 10,900 respondents were included in the final analytical sample. The range of eGFR values was chosen to exclude individuals with end-stage renal disease and to avoid estimating GFR in the highest range where precision is known to be low (165). The study population was followed for 4-5 years, from the baseline nurse health assessment in 2010-2013 to the seventh wave in 2015-2017.
Procedures for obtaining the anthropometric and health measures have been described in detail by McFall et al (137). The creatinine assay was standardized using IDMS. Grip strength was measured in the dominant hand in standing position with upper arm against the trunk and forearm at a right angle to the upper arm. The maximum reading from three attempts was recorded using a Smedley’s dynamometer. Standard deviation scores of grip strength were created within each sex and 10-year age groups starting at age 16 (ages above 86 years collapsed due to low numbers) and these were combined to form a standardised measure. Body weights above 130 kg were self-reported due to imprecision of the scale in that range.

Self-reported baseline diagnoses were identified in the wave that preceded the nurse health assessment; wave 2 for non-BHPS participants and the wave 3 interview for the BHPS sample. Chronic obstructive pulmonary disease (COPD) was identified as self-reported chronic bronchitis or emphysema. Baseline diabetes mellitus was defined as self-reported disease or a HbA1c value ≥48 mmol/mol. Data on smoking were not collected in wave 3 and therefore smoking status was determined at wave 2 for all participants. Smoking was classified as never regular smoker, ex-smoker or current smoker. Ethnicity was classified as white UK, Afro-Caribbean and other, consistent with the required information on Afro-Caribbean ethnicity required to estimate GFR using the CKD-EPI equation.

The outcome all-cause mortality was reported by the diseased individual’s household or identified through systematic enquiries in the event of non-contact in the following wave, otherwise participants were classified as alive.

Statistical analysis

Associations between eGFR and all-cause mortality were evaluated using logistic regression (due to the wave structure of the data) adjusting for age, sex, ethnicity, BMI, smoking and self-reported pre-existing diagnoses of CVD (ischemic heart disease or stroke), diabetes and hypertension. All analyses were adjusted for at the wave participants entered the study.

Linearity of the relationship between the log-odds of mortality and the continuous variables eGFR, age and BMI was assessed by applying the multivariable fractional polynomial (MFP) method (147). Selection of variables was pre-specified and no data-driven selection of variables was performed. The level of significance for transformation of eGFR was set to 5% and for transformation of potential confounding factors a 20% level of significance was applied. Age was modelled as linear in all models. BMI was modelled
as linear in the middle and highest strata of standardized grip strength. In the adjusted model before stratification and in the model for the lowest stratum of grip strength, BMI was modelled using a FP2 function with powers (1, 1). eGFR was modelled as non-linear in the adjusted model before stratification and in the adjusted model for the lowest stratum of grip strength using a FP1 function with power (-1). In sensitivity analyses not excluding individuals with eGFR outside the range of 15-120 ml/min/1.73 m² BSA, eGFR was modelled using a FP1 function with power (-0.5) in the adjusted model for the lowest stratum of grip strength. In sensitivity analyses adding additional adjustment for pre-existing cancer, pre-existing chronic obstructive pulmonary disease or pre-existing congestive heart failure, the same transformations as in the main analyses were selected.

OR for all-cause mortality were calculated at the median value of the following categories of eGFR: 30-44 (37.5), 45-60 (52.5), 60-89 (75), 90-104 (97.5) and 105-119 (112.5) ml/min/1.73 m² BSA using 90 ml/min/1.73 m² BSA as reference. No OR was calculated for the lowest eGFR category (15-29 ml/min/1.73 m² BSA) as the data were too sparse. The hypothesized interaction between eGFR and grip strength was evaluated by adding the main effects and the eGFR-grip strength interaction term to the adjusted model. The association between eGFR and all-cause mortality was further estimated by stratifying the adjusted model into thirds of the distribution of grip strength.
**Ethical considerations**

Paper I and paper III of this thesis are based on the Conscription Cohort which was assembled after ethical approval by the Regional Ethical Review Board of Uppsala Sweden (decision reference numbers 2009/306 and 2014/324), with the purpose to study associations between physical and psychological characteristics in adolescence and future chronic disease and mortality.

Paper II is based on the mGFR Cohort assembled locally in Örebro, specifically for the purpose of this project. Ethical approval was given by the Regional Ethical Review Board of Uppsala Sweden (decision reference number 2013/065).

Linkage between registers for papers I, II and III was carried out by Statistics Sweden and the National Board of Health and Welfare using personal identification numbers. In the data delivered to the researchers the personal identification numbers were removed in order to anonymise the data. Under these circumstances, the Ethical Review Board waived informed consent from the individual members of the study populations.

Paper IV utilizes data from the UKHLS made available to researchers by the UK Data Archive. No application to a Swedish Ethical Review Board was necessary as the data are made available for research purposes according to ethical permissions granted in the UK (Ethics Committee of the University of Essex and the National Research Ethics Service Oxfordshire REC A 10/H0604/2; 10/H0604/62; 10/H0604/70).

Data in all projects included in this thesis are stored securely by Örebro University and Region Örebro County by limited access and strict guidelines for use of the data.
Results

Paper I – Predictors in adolescence of ESRD in middle-aged men

Between January 1, 1985 and December 31, 2009, 534 incident cases of ESRD were identified and 5,127 men were selected as controls. The mean age at the initial conscription examination was 19 years, at start of follow-up 31 years and at the end of follow-up 55 years.

Table 7. Associations of characteristics in adolescence with ESRD

<table>
<thead>
<tr>
<th></th>
<th>Unadjusted</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (95% CI)</td>
<td>P</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td>ESR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.00</td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>2-6</td>
<td>1.13 (0.90-1.43)</td>
<td>0.3</td>
<td>1.10 (0.87-1.40)</td>
</tr>
<tr>
<td>7-10</td>
<td>1.53 (1.04-2.26)</td>
<td>0.03</td>
<td>1.39 (0.94-2.08)</td>
</tr>
<tr>
<td>11-14</td>
<td>2.03 (1.14-3.63)</td>
<td>0.02</td>
<td>1.93 (1.06-3.50)</td>
</tr>
<tr>
<td>≥15</td>
<td>2.46 (1.39-4.36)</td>
<td>0.002</td>
<td>2.07 (1.14-3.75)</td>
</tr>
<tr>
<td>Proteinuria</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>1.00</td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Positive</td>
<td>7.70 (4.02-14.75)</td>
<td>&lt;0.001</td>
<td>7.72 (3.94-15.14)</td>
</tr>
<tr>
<td>Blood pressure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;120/80</td>
<td>0.90 (0.67-1.19)</td>
<td>0.4</td>
<td>0.95 (0.71-1.26)</td>
</tr>
<tr>
<td>120-129/80-84</td>
<td>1.00</td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>130-139/85-90</td>
<td>1.21 (0.96-1.53)</td>
<td>0.1</td>
<td>1.19 (0.93-1.51)</td>
</tr>
<tr>
<td>140-159/90-99</td>
<td>1.66 (1.29-2.13)</td>
<td>&lt;0.001</td>
<td>1.45 (1.12-1.88)</td>
</tr>
<tr>
<td>≥160/≥100</td>
<td>5.16 (2.79-9.56)</td>
<td>&lt;0.001</td>
<td>3.97 (2.08-7.59)</td>
</tr>
<tr>
<td>BMI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;18.5</td>
<td>0.86 (0.64-1.17)</td>
<td>0.4</td>
<td>0.83 (0.61-1.15)</td>
</tr>
<tr>
<td>18.5 - &lt;25</td>
<td>1.00</td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>25 - &lt;30</td>
<td>2.63 (2.00-3.44)</td>
<td>&lt;0.001</td>
<td>2.36 (1.79-3.13)</td>
</tr>
<tr>
<td>≥30</td>
<td>4.73 (2.81-7.96)</td>
<td>&lt;0.001</td>
<td>3.53 (2.04-6.11)</td>
</tr>
</tbody>
</table>

*Adjusted for socioeconomic factors from the 1960 census (head of household’s occupation, household crowding) and characteristics in adolescence (ESR, EVF, proteinuria, blood pressure, BMI, physical working capacity and cognitive function score). ESR is expresses as mm/h; blood pressure (systolic/diastolic) as mmHg; BMI as kg/m².

The main findings in this case-control study nested within the Conscription Cohort, are that proteinuria is associated notably with future ESRD, with an adjusted OR of 7.72 (95% CI, 3.94-15.14; P<0.001). ESR has a dose-dependent association with ESRD with an adjusted OR of 2.07 (1.14-3.75; P=0.02) for ESR >15 mm/h. Hypertension is associated with future ESRD with an OR of 3.97 (2.08-7.59; P<0.001) for grade 2 hypertension and higher. Elevated BMI is associated statistically significantly with increased ESRD risk with an OR of 3.53 (2.04-6.11; P<0.001) for BMI ≥30.
compared with 18.5-<25 kg/m². These associations are independent of each other and extend up to 40 years into adult life from late adolescence. OR for the associations with ESRD for the evaluated markers of risk are given in table 7.

**Paper II – Measured glomerular filtration rate does not improve prediction of mortality by cystatin C and creatinine**

The mean duration of follow-up was 4.71 years and during that time 312 patients died, equivalent to a mortality rate of 5.73 individuals/100 person-years among the 1,157 individuals in the cohort. The high mortality rate reflects that the study population were patients referred for measurement of GFR rather than a general population sample.

![Figure 3. Functional form of the associations with all-cause mortality for cystatin C, creatinine and mGFR in Cox regression models, including models where these measures are combined. The histograms illustrate the distribution of each marker.](image)

---

50  PER-OLA SUNDIN  A life-course approach to chronic kidney disease
The main findings of this cohort study are that after adjustment for mGFR, a higher cystatin C concentration remains associated with increased mortality but for creatinine the association is reversed so that a lower creatinine concentration is associated with increased mortality. The functional form of the associations with mortality for markers of GFR, including changes in functional form when these measures are combined are illustrated in figure 3. When nested models are compared, to reduce the potential influence of multicollinearity, the independence of the associations is supported. Among models combining the markers of GFR, adjusted for demographic factors and comorbidity, cystatin C and creatinine combined explain the largest proportion of variance in associations with mortality risk ($R^2=0.61$) and addition of mGFR does not improve this model. Both cystatin C and creatinine increases the discriminative power (category free NRI) of the adjusted model which includes mGFR but for creatinine the improvement is limited to increased specificity. Based on these results, the model combining cystatin C$^{-1}$ and creatinine$^{-1}$ while not including mGFR was selected as the final model.

The distribution of serum creatinine$^{-1}$ and serum cystatin C$^{-1}$ and plots of the HR for all-cause mortality in the final fully adjusted model combining cystatin C and creatinine are displayed in figure 4.

**Paper III – Hospital admission with pneumonia and subsequent persistent risk of chronic kidney disease**

Over a median of 36.7 (interquartile range 35.3-37.9) years of follow-up, 5,822 men were discharged from hospital diagnosed with pneumonia and without a diagnosis of kidney disease. Excluding the first year, 1.73% ($n=101$) of men with pneumonia were later diagnosed with CKD compared to 1.2% ($n=2,749$) of non-exposed. During follow-up 19.9% of the exposed men with CKD progressed to ESRD.

The main findings of this cohort study using the Conscription Cohort are a high magnitude association between hospital admission with pneumonia and future incident CKD which attenuates with time but remains statistically significant more than 15 years later. HR (and 95% CI) for subsequent CKD and ESRD after excluding the first year after pneumonia, are: 5.20 (3.91 to 6.93) and 5.72 (3.14 to 10.41) in the following five years, adjusted for potential confounding factors measured at conscription assessment in late adolescence identified in paper I. Sensitivity analyses do not indicate any notable change in the associations, apart from additional adjustment
Figure 4. Plots of hazard ratios for all-cause mortality in the final fully adjusted Cox regression model. Hazard ratios are compared to reference points within the normal range using 1.08 (cystatin C 0.925 mg/L) for cystatin C⁻¹ and 0.013 (creatinine 75 µmol/L) for creatinine⁻¹. Dashed lines indicate 95% CI. The histograms illustrate the distribution of each marker. X axes are labeled to illustrate both reciprocals and actual values.
for CVD, diabetes and hyperlipidaemia as time-dependent covariates, which attenuates the magnitude of the HR moderately to levels comparable to the results in the matched hospital cohort (table 8).

In the nested matched hospital cohort additionally adjusted for CVD and diabetes, the HR for CKD is 2.85 (1.84-4.11) during the first five years and 1.75 (1.09-2.82) during the following 10 years (table 8).

**Table 8.** Associations of pneumonia with chronic kidney disease and end-stage renal disease during time periods following hospital admission with pneumonia

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Unadjusted HR (95% CI)</th>
<th>Adjusted model HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chronic kidney disease</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main analysis&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 5 years</td>
<td>5.63 (4.23 to 7.49)</td>
<td>5.20 (3.91 to 6.93)</td>
</tr>
<tr>
<td>&gt;5-≤15 years</td>
<td>2.90 (2.07 to 4.07)</td>
<td>2.80 (1.99 to 3.93)</td>
</tr>
<tr>
<td>&gt;15 years</td>
<td>2.17 (1.38 to 3.42)</td>
<td>2.13 (1.35 to 3.35)</td>
</tr>
<tr>
<td>Comorbidity&lt;sup&gt;ab&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 5 years</td>
<td>-</td>
<td>2.44 (1.82-3.25)</td>
</tr>
<tr>
<td>&gt;5-≤15 years</td>
<td>-</td>
<td>1.69 (1.20-2.38)</td>
</tr>
<tr>
<td>&gt;15 years</td>
<td>-</td>
<td>1.66 (1.06-2.62)</td>
</tr>
<tr>
<td>Matched cohort&lt;sup&gt;ac&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 5 years</td>
<td>2.81 (1.97-4.02)</td>
<td>2.85 (1.84 to 4.41)</td>
</tr>
<tr>
<td>&gt;5-≤15 years</td>
<td>2.13 (1.42 to 3.20)</td>
<td>1.75 (1.09 to 2.82)</td>
</tr>
<tr>
<td><strong>End-stage renal disease</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main analysis&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 5 years</td>
<td>6.22 (3.42 to 11.32)</td>
<td>5.72 (3.14 to 10.41)</td>
</tr>
<tr>
<td>&gt;5 years</td>
<td>2.20 (1.17 to 4.12)</td>
<td>2.15 (1.15 to 4.03)</td>
</tr>
</tbody>
</table>

Follow-up time starting one year after pneumonia. <sup>a</sup>Adjusted for characteristics in adolescence (BMI, blood pressure, physical working capacity, cognitive function, ESR, EVF) and childhood (household crowding and head of household’s occupation).  
<sup>b</sup>Additionally adjusted for CVD, diabetes and hyperlipidaemia as time varying covariates.  
<sup>c</sup>Additionally adjusted for CVD and diabetes.
Paper IV – Grip strength modifies the association between estimated glomerular filtration rate and all-cause mortality

During 4-5 years of follow-up, mortality was recorded for 2.48% (270/10,900) of the study population. Compared to the group with high grip strength, individuals with low grip strength have a lower serum creatinine and a lower BMI; are more likely to belong to ethnic groups other than white UK or Afro-Caribbean and more often report pre-existing diabetes and CVD.

The main findings of this cohort study utilizing a subsample of the UKHLS general population sample, are effect modification by grip strength for the association between eGFR and increased mortality risk indicating an increased vulnerability to the adverse effects of reduced eGFR. Adjusted OR for all-cause mortality calculated at the median value (with range) of the following categories of eGFR: 37.5 (30-44), 52.5 (45-60), 75 (60-89), 97.5 (90-104) and 112.5 (105-119) ml/min/1.73 m² BSA using 90 ml/min/1.73 m² BSA as reference are presented in table 9. The corresponding unadjusted OR (and 95% CI) decrease linearly from 12.8 (8.90-16.67) at eGFR 37.5 ml/min/1.73 m² BSA to 0.34 (0.30-0.73) at eGFR 112.5 ml/min/1.73 m² BSA compared with eGFR 90 ml/min/1.73 m² BSA. In the adjusted model there is a statistically significant U-shaped association between eGFR and all-cause mortality. Standardized grip strength is not correlated with eGFR or any transformations of eGFR used in the models (Pearson’s correlation coefficients 0.04-0.01). However, after further adjustment for grip strength, the U-shaped association is transformed into a more linear pattern where lower eGFR is statistically significantly associated with increased mortality risk; OR decrease from 1.70 (1.18-2.46) at eGFR 37.5 ml/min/1.73 m² BSA to 0.93 (0.88-0.98) at eGFR 112.5 ml/min/1.73 m² BSA compared with eGFR 90 ml/min/1.73 m² BSA. This is explained by a statistically significant multiplicative interaction between grip strength and eGFR (p=0.04). After stratification into thirds of the distribution of standardized grip strength, only the lowest stratum has statistically significant associations between lower eGFR and increased mortality risk, while in the higher strata there is no raised risk (table 9).

Additional adjustment for self-reported COPD, cancer or congestive heart failure does not have any notable effect on the magnitude of the associations between eGFR and all-cause mortality in the stratified models. Repeating the analyses also including individuals with eGFR outside the range of 15-120 ml/min/1.73 m² BSA, did not change the results notably.
Table 9. Associations of eGFR with all-cause mortality

<table>
<thead>
<tr>
<th>eGFR</th>
<th>Total (n= 10,900)</th>
<th>Lowest third (n=3,637)</th>
<th>Middle third (n=3,660)</th>
<th>Highest third (n=3,603)</th>
</tr>
</thead>
<tbody>
<tr>
<td>37.5</td>
<td>1.68 (1.09-2.61)</td>
<td>2.24 (1.39-3.61)</td>
<td>0.94 (0.38-2.29)</td>
<td>0.83 (0.30-2.34)</td>
</tr>
<tr>
<td>52.5</td>
<td>1.07 (0.77-1.50)</td>
<td>1.50 (1.18-1.91)</td>
<td>0.96 (0.51-1.81)</td>
<td>0.88 (0.42-1.83)</td>
</tr>
<tr>
<td>75.0</td>
<td>0.86 (0.71-1.04)</td>
<td>1.12 (1.05-1.20)</td>
<td>0.98 (0.76-1.27)</td>
<td>0.95 (0.71-1.27)</td>
</tr>
<tr>
<td>90.0</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>97.5</td>
<td>1.18 (1.03-1.35)</td>
<td>0.96 (0.93-0.98)</td>
<td>1.01 (0.89-1.15)</td>
<td>1.03 (0.89-1.19)</td>
</tr>
<tr>
<td>112.5</td>
<td>1.95 (1.20-3.17)</td>
<td>0.89 (0.83-0.95)</td>
<td>1.03 (0.70-1.50)</td>
<td>1.08 (0.70-1.68)</td>
</tr>
</tbody>
</table>

Adjusted for age; sex; ethnicity; BMI; smoking and self-reported pre-existing diagnoses of CVD, diabetes and hypertension. *Stratified into thirds of the distribution of grip strength. Grip strength is standardized by age and sex.
Discussion

The two main themes of this thesis are identification of risks for future CKD and predicting consequences of established CKD in terms of all-cause mortality, from a life-course perspective. This discussion section will focus on contextualising our findings and possible implications.

Findings and implications

Risks for CKD from a life-course perspective

Few studies have investigated CKD risk from a life-course perspective. The hereditary predisposition to develop CKD constitutes a risk that operates across generations (62-64). Exposures during gestation resulting in prematurity or low birth weight, can result in a reduced number of nephrons and increased risk for hypertension, thus increasing CKD risk throughout the life course (73). Obesity in childhood and adolescence increase CKD risk in adulthood (74, 75). The remaining well-established risk factors for CKD have mainly been investigated in adulthood for largely contemporaneous associations in the same stage of the life course.

The findings in paper I are consistent with earlier studies finding associations of proteinuria, hypertension, elevated BMI and markers of inflammation with future CKD or ESRD. The accumulation of these risk factors up to late adolescence may reflect earlier exposures possibly beginning in utero. Apart from previous studies linking elevated BMI in childhood and adolescence with adult CKD, our findings provide novel evidence indicating that associations with increased risk of ESRD extend from childhood and adolescence into middle-age, up to 40 years after initial assessment.

The results in paper III suggests that hospital admission with pneumonia in adulthood is associated, not only with acute onset kidney disease, but also a persistent increased risk of subsequent CKD. The highest magnitude risk is during the years immediately following pneumonia, and although attenuated over time, the raised risk of CKD remained statistically significant for more than 15 years. This risk is independent of the markers of risk at conscription assessment reported in paper I and therefore seems to indicate the adult origin of another pathway to CKD.

It is noteworthy that our novel findings regarding CKD risks in both papers I and III are consistent with earlier findings for CVD risks. An origin of CVD in childhood and adolescence conferring increased risk in later stages of life is well established (71), and multiple studies have identified
associations between factors in childhood and adolescence, including higher ESR, hypertension and higher BMI in Swedish adolescents, and adult CVD (80, 166-172). Our findings in paper III of CKD following pneumonia, are also consistent with earlier studies demonstrating an increased risk of CVD persisting for more than 10 years following pneumonia (86). This can be understood from the similarities of the pathophysiological processes of CKD and CVD which in part are parallel processes, both associated with inflammation (57-60). Accordingly, CKD and CVD share several conventional risk factors and display a bi-directional association, each being associated with a high magnitude risk of developing the other condition (47, 48).

In paper I men with a CVD diagnosis at conscription assessment were excluded when the risk sets were constructed but information on possible later CVD preceding CKD and later ESRD was not included in the analysis. Thus, possible mediation of the associations between markers in adolescence and later ESRD through CVD cannot be determined. However, this does not invalidate the potential usefulness of these markers for identification of future elevated ESRD risk. In paper III, sensitivity analyses, including adjustment for CVD as a time-dependent covariate and the construction of a nested matched hospital cohort with adjustment for CVD, investigated possible confounding from CVD. These demonstrated that although CVD was associated with a higher magnitude risk of CKD following pneumonia, a history of inpatient care for pneumonia identified a statistically significant independent risk.

Our findings in paper I and paper III both highlight the long natural history of CKD and ESRD, and emphasize the need to adopt a life-course approach to understand and prevent CKD. When incident cases of CKD are identified in adulthood, risk factors for CKD present at the time and more recent medical history are obvious factors to be considered when the clinical assessment of possible aetiology is made. This translates into targets for secondary prevention but also into current views on screening for CKD and primary prevention of CKD. However, early life events and the accumulation of risk during childhood and adolescence as well as infectious events across the life-course, may be important factors in the pathway to CKD that should be recognised. Our results may contribute to the accumulation of evidence needed to identify target populations for future screening for CKD in addition to the high risk populations screened today. The results in paper I also contribute to our understanding of the possible impact on adult health.
of childhood factors and the importance of preventive measures in childhood and adolescence.

**Consequences of CKD**

In paper II we demonstrate how creatinine and cystatin C both are associated with mortality beyond what is indicated by mGFR and that mGFR does not offer a more precise mortality risk assessment than cystatin C and creatinine combined. This is clearly demonstrated when the functional form of the associations between the reciprocals of these filtration markers (due to the invers relation to GFR) alone or combined are examined (figure 3) and when the performance of nested models is compared. Associations between endogenous filtration markers and increased mortality risk independent of mGFR suggest the existence of factors associated both with mortality risk and the production rate, sieving coefficient or extra-renal clearance of the endogenous filtration marker.

In the case of creatinine, low creatinine production, mainly seen in individuals with low muscle mass, is an established powerful predictor of mortality. The J-shaped association for creatinine⁻¹ where both high values (indicating low muscle mass/low creatinine production or hyperfiltration) and low values (low GFR) are associated with increased mortality disappears when another measure of filtration is introduced in the model, leaving creatinine as primarily a marker of creatinine production (figure 3). This attenuates the association between higher creatinine concentration and increased mortality when a single marker of GFR is used. However, when Cystatin C and creatinine were combined in our models, associations between low creatinine production and mortality is likely to have contributed to the models’ ability to identify increased mortality risk. Previously, timed urinary samples to quantify the amount of creatinine excreted have been required to utilize creatinine’s non-GFR associations with mortality, apart from in dialysis patients with negligible GFR (117-119, 173, 174).

Associations of cystatin C with mortality independent of mGFR has been suggested in earlier studies but interpretation has been complicated by the high degree of collinearity between cystatin C and mGFR (114-116). Our findings of an association between cystatin C and mortality independent of mGFR which is not explained by effects of multicollinearity, raises the question of which mechanisms explain this association.

The implications of our results in paper II are that if the purpose is to evaluate the increased mortality risk associated with reduced GFR per se, mGFR is required, but if the purpose is to optimally predict mortality risk
The combination of cystatin C and creatinine is best suited. This is further discussed in an invited ‘in focus’ commentary accompanying our paper in Nephrology Dialysis Transplantation (107).

The Shrunken Pore Syndrome

The mechanisms explaining our findings of an association with increased mortality risk for higher cystatin C independent from mGFR, has not been established. Associations between cystatin C and inflammation might explain these findings and higher cystatin C values have been shown to be correlated with higher C-reactive protein levels indicating inflammation (33, 175). This suggests increased production of cystatin C in inflammatory states. However, studies on patients undergoing elective surgery provoking an inflammatory response with elevated C-reactive protein levels has not shown concomitant increased levels of cystatin C (37, 176). Thus, the association between C-reactive protein and cystatin C is not causal. Rather, the association might be explained by inflammation promoting kidney damage which reduces GFR and thereby increases cystatin C levels (37).

Another possibility is that a pathophysiological state associated with increased mortality, reduced the sieving coefficient for cystatin C (lower permeability for cystatin C in the glomerular filtration barrier) but did not affect the filtration of creatinine to the same extent. Creatinine is a small molecule with a molecular weight of 113 Da, while the larger cystatin C molecule has a molecular weight of 13.3 kDa. In 2015, Grubb et al. described such a pathophysiological state and named the associated syndrome the shrunken pore syndrome (177). The concept of filtration quality (differences in the composition of the glomerular filtrate) as a tool in diagnosing kidney disease, beyond quantifying albuminuria and measuring the quantity of filtrate (GFR) had already been suggested by Grubb et al. in 2007 (178).

Current evidence and the evolution of research regarding the shrunken pore syndrome are outlined here briefly based on a recent review in Swedish by Grubb (175). When cystatin C was first introduced as a filtration marker it was soon recognised that the glomerular filtration of creatinine and cystatin C differ during the course of pregnancy. The serum concentration of creatinine stayed fairly constant while the cystatin C concentration increased to reach a maximum in the end of the third trimester (74, 179-181). This translates into a progressively lower ratio between eGFR estimated from cystatin C and eGFR estimated from creatinine, starting close to 1 which is normal in healthy non-pregnant women, and reaching approximately 0.6 at the end of pregnancy (177, 180, 181). Preeclampsia was
shown to be associated with even lower eGFR cystatin C / eGFR creatinine ratios (74, 179-181).

Grubb et al identified in 2015 the same situation with an eGFR cystatin C / eGFR creatinine ratio below 0.6 in 8.2% of a population of men and non-pregnant women. They further showed that not only the cystatin C concentration, but also the concentrations of other molecules with a molecular weight between 5 and 40 kDa were increased in relation to the creatinine concentration in this pathophysiological state. This selective effect on the concentrations of molecules which depend on normal pores in the glomerular filtration barrier for their excretion in the urine, while not affecting the excretion of small molecules with a molecular weight below 0.2 kDa including creatinine and water, led to the suggestion of the *shrunken pore syndrome* hypothesis (177).

Molecules used as exogenous filtration markers including iohexol (821 Da) are in general of a molecular weight below 5 kDa and are therefore not sensitive to the changes in filtration described by the shrunken pore syndrome.

Dardashti et al. found a prevalence of the shrunken pore syndrome of 5.6% using the same definition (eGFR cystatin C / eGFR creatinine <0.6) in a population treated with coronary artery bypass grafting. For individuals with the syndrome, the five year survival was 65%, compared with 90% in patients without the syndrome (182). Even in people with a normal eGFR, the presence of the shrunken pore syndrome was associated with an increased mortality risk. Later studies in different populations have consistently found associations between the shrunken pore syndrome and increased mortality risk, mainly due to cardiovascular morbidity (183-187).

A recent study of the proteome of patients with and without the shrunken pore syndrome and with and without reduced mGFR, found an accumulation of several atherosclerosis promoting proteins specifically in individuals with the shrunken pore syndrome (188). Thus, besides the possibility of the shrunken pore syndrome being the result of already established microvascular disease, the accumulation of atherosclerosis promoting proteins as a result of the altered filtration, may explain associations of the shrunken pore syndrome with increased mortality and morbidity.

The accumulating evidence of a change in filtration quality, detectable by comparing eGFR from cystatin C and eGFR from creatinine, associated with increased mortality and CVD, favours an approach were both endogenous filtration markers are determined when assessments of future risk are to be made. This can be achieved by adopting a model combining these two
measures as in paper II. However, in a clinical setting, similar ability to identify increased mortality risk might be accomplished by treating the mean of eGFR from cystatin C and eGFR from creatinine as a measure of GFR indicating risks associated with reduced GFR (filtration quantity) and the ratio between eGFR from cystatin C and eGFR from creatinine in addition to a measure of albuminuria to identify risks associated with altered filtration quality. This however needs to be evaluated in future research.

**Differential associations of eGFR with all-cause mortality when stratified by grip strength**

When assessing individual increased mortality risk associated with a given reduced level of eGFR, multiple factors need to be considered, including the following. First, the precision of eGFR based on a single endogenous filtration marker is quite low. In a recent opinion paper by Porrini et al., this was summarized stating that 10-40% of eGFR differs by more than 30% from mGFR and that eGFR misclassify 30-60% of subjects for CKD stage. Second, risk estimates provided for stages of CKD relate to chronic disease where the reduced eGFR has been recorded for at least three months. Many studies investigating adverse outcomes from CKD, including paper II and paper IV of this thesis, rely on single measurements of markers of kidney function and this tends to overestimate CKD prevalence (12). Third, the relative increase in mortality risk associated with lower eGFR is attenuated with increasing age (189). This relates to the fact that although GFR is known to decline with increasing age, a uniform threshold for CKD at an eGFR of 60 ml/min/1.73 m² BSA is applied in the KDIGO guidelines (table 1) (5). Thus, the threshold for CKD is not determined by the distribution of eGFR at a given age in a healthy population but by the level were statistically significant increased risk is observed (190). The results of paper II add to this complexity by providing further evidence for association with mortality for creatinine and cystatin C beyond their reflection of GFR.

In paper IV we investigated possible effect modification by grip strength for the association between eGFR estimated from serum creatinine and all-cause mortality. Grip strength correlates moderately with muscle mass which is the most important determinant of the production rate of creatinine. The research question arose from the results in paper II for the functional form of the association between creatinine⁻¹ and all-cause mortality in the different models. If muscle mass confounds the association, adjustment for muscle mass would potentially change the J-shaped association into a more linear association where higher creatinine (lower creatinine⁻¹,
lower eGFR) is associated with increased mortality. However, both reduced muscle mass and reduced muscle strength are components of the closely related clinical syndromes of sarcopenia and frailty which entail an increased vulnerability to endogenous and exogenous stressors (191). This implies a possible more complex interplay between muscle mass and muscle strength with eGFR for their associations with mortality, where sarcopenic/frail individuals would be more vulnerable to the adverse effects of reduced GFR. Given the changes in metabolic, inflammatory and hormonal systems that have been associated with sarcopenia and low muscle strength, possibly important for the adverse effects of reduced GFR, this hypothesized differential vulnerability is in our opinion also biologically plausible (118, 192, 193).

In paper IV we used muscle strength measured as hand grip strength as a marker of muscle function associated with muscle mass. Grip strength was standardised by age and sex. Low grip strength was associated with increased mortality, but not with eGFR or any transformations of eGFR used in the models. Thus, grip strength did not confound the association between low eGFR and increased mortality risk. The hypothesized multiplicative interaction between grip strength and eGFR for mortality risk was statistically significant, and stratification of the model into thirds of the distribution of grip strength illustrated how only individuals with low grip strength displayed an increased mortality risk with lower eGFR after adjustment for the potential confounding factors.

These novel findings add another dimension of complexity to the assessment of mortality risk associated with low eGFR from creatinine. They imply that the increased all-cause mortality risk associated with moderately reduced eGFR over the next 4-5 years is not notably raised for individuals in the general population whose grip strength is not reduced compared with individuals of the same age and sex. Thus, testing grip strength may improve stratification of all-cause mortality risk based on eGFR.

Methodological considerations
In addition to the methodological considerations, including strengths and limitations, already discussed throughout the methods and results section as well as in the individual papers, I would like to make a few further comments on subjects shared by several of the papers in this thesis.
**Routinely collected health data**

The unique personal identification number issued to all residents of Sweden allows almost complete follow-up of each individual throughout the life-course, gathering information left in Swedish national population and health registers (129). This provides excellent conditions for register-based research with the possibility to extend follow-up over decades, thereby spanning different stages of the life-course. Although, paper I and paper III were retrospectively defined studies, the routinely collected health-data used was collected prospectively.

Other benefits of using routinely collected health data are entire population coverage which allow studies of relatively rare outcomes (ESRD in paper I); the avoidance of non-response, attrition and reporting bias frequent in surveys; time-efficient and cost-effective compared with prospective data collection (194). However, there are also several potential limitations including event based data where individuals without an event such as a hospital stay will have no records (surveillance bias for CKD in paper III); possibly limited information on important potential confounding and risk factors (smoking and baseline GFR in paper I) (194).

In my opinion there is also a risk that research questions are adapted to the data available, but still I believe that it would be unethical not to make efforts to extract information that might improve healthcare from the vast datasets that are recorded for primarily administrative purposes.

**Generalizability of results from men to women**

The Conscription Cohort used in paper I and paper III include routinely collected health data from the Swedish Military Conscription Register. Since military conscription examinations were compulsory only for males until relatively recently, there is no information on women in the same birth cohorts. This is a major limitation since there are indications of differences in the underlying pathophysiology of CKD between men and women (195, 196). CKD is more common among women, GFR declines faster in men compared with women and mortality is higher for men compared with women during all stages of pre-dialysis CKD while this difference disappears at start of renal replacement therapy (196). This limits the generalizability of the results in paper I and paper III to women.
Conclusions

CKD and ESRD have a long natural history where novel insights into risk factors for CKD may be obtained by adopting a life-course perspective.

- Proteinuria, hypertension, higher ESR and higher BMI in late adolescence are associated with ESRD in middle-aged men up to 40 years after the initial assessment.
- Hospital admission with pneumonia in adulthood is associated with a persistent increased risk for subsequent CKD, with the highest magnitude risk in the years immediately after infection, then attenuating over time, but the raised risk remains statistically significant for more than 15 years.

When assessing individual increased all-cause mortality risk associated with endogenous filtration markers some novel aspects may be considered.

- If the purpose is to evaluate the risk associated with reduced GFR per se, mGFR is required but if the purpose is to optimally predict mortality risk the combination of cystatin C and creatinine is best suited.
- Low grip strength may be a readily available measure which identifies more accurately individuals with low eGFR from serum creatinine at increased risk of all-cause mortality.
Future perspectives

Considering the results of paper I indicating associations between markers of health and development in late adolescence and future ESRD risk up to 40 years later, additional potential markers of increased CKD risk in childhood and adolescence, should be investigated and the specific biological pathways identified to facilitate prevention. We plan to analyse possible associations of grip strength measured at the conscription assessment in late adolescence with CKD and ESRD in adulthood using the Conscription Cohort.

Based on the results in paper II and the expanding research on the hypothesized pathophysiological state described as the shrunken pore syndrome, we plan to study prediction of cause-specific mortality from the eGFR cystatin C / eGFR creatinine ratio, and its independence from mGFR in the mGFR Cohort. The total cohort comprises over 2,000 patients but the older sub-cohorts, with cystatin C measurements before standardization of cystatin C analyses was available, were not included in paper II of this thesis. We plan to use published eGFR equations developed in the same sub-cohorts to calculate eGFR from serum cystatin C separately in each sub-cohort and then analyse the whole cohort using multilevel modelling and Cox regression.

In paper III we found associations between hospital admission with pneumonia and future persistent CKD risk, still statistically significant over 15 years later. Information on kidney function other than from diagnostic codes in the NPR was not available and future prospective studies or retrospective studies utilizing routinely collected health data, are needed where estimates of GFR are included at baseline and during follow-up.

Future studies are needed to confirm the observed effect modification by grip strength for the association between low eGFR and increased all-cause mortality risk indicated by our results in paper IV. Studies should then include both eGFR and mGFR, as well as measures of muscle strength and muscle mass.
Acknowledgements

I would like to express my sincere gratitude to all of you who in different ways have contributed to this work. In particular, I would like to thank:

Scott Montgomery, my main supervisor and role model in research, for mastering the ability to welcome any question and disturbance no matter how busy you are. Knowing that you always will guide me to a solution to any problem I might face in research has been so important to me during these years. Your analytical brilliance and excellence in scientific writing has been a true inspiration.

Katja Fall, my co-supervisor, for always being encouraging and supportive. Your ability to make any constructive feedback feel like a push forward is a true gift.

Per Sjöström, my co-supervisor, for being such an inspiration. Thank you for inviting me to your research group and for sharing so generously your expertise in measuring kidney function. Your kindness and patience have been heartwarming. After getting to know you Per, I finally understood the meaning of ‘successful ageing’.

Ruzan Udumyan, co-author of all four papers, for all I have learnt from you about statistics and STATA, and for you making learning so enjoyable. Thank you for bringing such quality and complexity to the analysis of data in paper I and paper II, your contribution has been invaluable.

Anders Grubb, co-author of paper II, for sharing data and expertise so generously. Your continued support has been much appreciated.

My co-authors Ian Jones, Lovisa Olsson and Veronica Lindström, thank you for stimulating collaborations and valuable contributions.

All my colleagues at Clinical Epidemiology and Biostatistics, including prior and current PhD-students, for making me feel welcome and for creating such a friendly and inspiring atmosphere. Thank you Cecilia Bergh and Snieguole Vingeliene, for making sharing an office a pure delight.
Mikael Rizell, my role model in medicine. Knowing what kind of physician I want to become has been so important for the choices I have made. Thank you for always encouraging me to get into research and for being the most loyal and considerate colleague I have ever came across.

All my friends at Närjsjukvårdssteam Örebro, for making my clinical work a pure joy and for loyal support whenever my workload has felt too heavy.

I acknowledge Elsevier, the National Kidney Foundation, Oxford University Press and Dove Press for granting permission to reprint the individual papers in this thesis.

Region Örebro County for continued financial support, including grants awarded me from funding provided by the Swedish state under the agreement between the Swedish government and the county councils, the ALF-agreement.

The UK Economic and Social Research Council (ESRC) for grants to the International Centre for Life Course Studies (grants RES-596-28-0001 and ES/J019119/1) making this research possible.

Finally and most importantly, my beloved wife Anette and our children Felix, Julia, Casper and Engla for putting up with me during these years. I am so proud to be the husband of Anette and the father of our children. Thank you for all the joy and love you give.
Sammanfattning på svenska

Kronisk njursjukdom med nedsatt njurfunktion (glomerulär filtrationshastighet, GFR) är ett växande folkhälsoproblem med en uppskattad prevalens på 8-16% av befolkningen vilket beräknas förklara cirka 1,23 miljoner dödsfall årligen globalt.

Framgångsrikt förebyggande arbete mot kronisk njursjukdom förutsätter kunskap om vilka faktorer som initierar utvecklingen mot kronisk njursjukdom, när i livet dessa processer uppstår och hur de kan spänna över livets olika faser. Delarbete I och III i denna avhandling studerar riskfaktorer för kronisk njursjukdom i ett sådant livsförloppsperspektiv i en stor födelsekohort av Svenska män födda 1952-1956 som genomgått obligatorisk mönstring för militärtjänst och sedan följts genom hälsodata i nationella register.

I delarbete I undersökte vi sambandet mellan olika markörer för hälsa bestämda vid mönstring i sena tonåren och risken för att upp till 40 år senare utveckla terminal njursvikt. Såväl fynd av äggvita i urinen, högre sänka (erytrocyternas sänkningsreaktion), högre blodtryck och högt BMI (body mass index) var oberoende av varandra associerade med en kraftigt förhöjd risk. I delarbete 3 undersökte vi om män som vårdats på sjukhus med lunginflammation i vuxen ålder löper en förhöjd risk att senare diagnostiseras med kronisk njursjukdom oberoende av de riskmarkörer för terminal njursvikt vid mönstring som påvisats i delarbete I. Vi fann en kraftigt ökad risk för kronisk njursjukdom som trots att den avklingar med tiden fortfarande är statistiskt signifikant mer än 15 år senare.

Delarbete II och IV i denna avhandling studerar hur olika markörer för nedsatt GFR kan förutsäga den ökad dödlighet som följer med nedsatt njurfunktion. Delarbete II baseras på en konsekutiv serie av patienter som genomgått en noggrann bestämning av GFR vid Universitetssjukhuset Örebro (mGFR) och sedan följts genom hälsodata i Svenska nationella register. För enklare skattningar av GFR i klinisk vardag (estimerat GFR, eGFR) används framför allt serumkoncentrationen av kreatinin och cystatin C vilka dock påverkas även av andra faktorer än GFR. Dessa icke-GFR faktorer skulle i sig kunna vara relaterade till ökad dödlighet. I vår analys hade såväl kreatinin som cystatin C associationer till ökad dödlighet som inte helt kan förklaras av mGFR. När cystatin C och kreatinin kombinerades i samma modell kunde risk för ökad dödlighet kopplad till såväl nedsatt GFR som andra icke-GFR faktorer identifieras. Information om mGFR förbättrade inte modellen. Detta talar för att om risken för ökad dödlighet kopplad till nedsatt GFR ska bestämmas krävs en mer avancerad bestämning av GFR (mGFR),
men om avsikten är att få den mest precisa förutsägelsen av ökad dödlighet bör de enklare markörerna kreatinin och cystatin C kombineras.

Delarbete IV analyserar data från United Kingdom Household Longitudinal Survey (UKHLS) vilken följer medlemmar av cirka 40 000 hushåll i Storbritannien. Den viktigaste faktor utöver GFR som påverkar serumkonzentrationen av kreatinin är produktionshastigheten i muskulaturen. För de deltagare där biomarkörer bestämts i blodprover, studerade vi om risken för ökad dödlighet kopplad till lågt eGFR beräknat från serum kreatinin, skiljer sig för personer med olika handstyrka. Vi fann att handstyrka modifierar effekten av eGFR på risken för ökad dödlighet så att detta samband var statistiskt signifikant endast i den tredjedel av deltagarna med lägst handstyrka. Detta innebär att test av handstyrka skulle kunna vara ett sätt att förfina förutsägelser om framtida dödlighet baserade på reducerat eGFR beräknat med serum kreatinin.

Sammanfattningsvis understryker resultaten av våra arbeten vikten av att ha ett livsförloppsperspektiv då riskfaktorer för kronisk njursjukdom studeras, då dessa samband kan verka över långa tider och spänna över livets olika faser. Ytterligare slutsatser är att mer avancerade mätningar av GFR inte ger mer information om framtida dödlighet än cystatin C och kreatinin kombinerat och att känsligheten för de negativa effekterna på överlevnaden av ett reducerat eGFR kan skilja sig mellan olika individer, så som illustre-rats för personer med olika handstyrka.
References


161. Pearson K. X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science. 1900;50(302):157-75.


PER-OLO SUNDIN  A life-course approach to chronic kidney disease  89
Publications in the series
Örebro Studies in Medicine


35. Söderqvist, Fredrik (2009). Health symptoms and potential effects on the blood-brain and blood-cerebrospinal fluid barriers associated with use of wireless telephones.
41. Gustafsson, Sanna Aila (2010). The importance of being thin – Perceived expectations from self and others and the effect on self-evaluation in girls with disordered eating.
42. Johansson, Bengt (2010). Long-term outcome research on PDR brachytherapy with focus on breast, base of tongue and lip cancer.
43. Tina, Elisabet (2010). Biological markers in breast cancer and acute leukaemia with focus on drug resistance.
46. de Leon, Alex (2010). Effects of Anesthesia on Esophageal Sphincters in Obese Patients.


52. Loiske, Karin (2011). Echocardiographic measurements of the heart. With focus on the right ventricle.


64. Nordin Olsson, Inger (2012). Rational drug treatment in the elderly: "To treat or not to treat".


67. Thuresson, Marie (2012). The Initial Phase of an Acute Coronary Syndrome. Symptoms, patients’ response to symptoms and opportunity to reduce time to seek care and to increase ambulance use.


75. Gustavsson, Anders (2012): Therapy in Inflammatory Bowel Disease.
83. Lönn, Johanna (2013): The role of periodontitis and hepatocyte growth factor in systemic inflammation.
96. Sundh, Josefin (2013): Quality of life, mortality and exacerbations in COPD.
98. Palmetun Ekbäck, Maria (2013): Hirsutism and Quality of Life with Aspects on Social Support, Anxiety and Depression.


109. Törös, Bianca (2014): Genome-based characterization of Neisseria meningitidis with focus on the emergent serogroup Y disease


120. Pelto-Piri, Veikko (2015): Ethical considerations in psychiatric inpatient care. The ethical landscape in everyday practice as described by staff.


139. Elwin Marie (2016): *Description and measurement of sensory symptoms in autism spectrum.*

140. Östlund Lagerström, Lina (2016): "The gut matters" - an interdisciplinary approach to health and gut function in older adults.


157. Olsson, Emma (2017): Promoting Health in Premature Infants – with special focus on skin-to-skin contact and development of valid pain assessment.


177. Christos Karefylakis (2018): Vitamin D and its role in obesity and other associated conditions.


188. Röckert Tjernberg, Anna (2019): *Celiac Disease and Infections.*


