Epidemiological aspects of health and work ability in the Swedish workforce

Adnan Noor Baloch

School of Public Health and Community Medicine
Institute of Medicine
Sahlgrenska Academy, University of Gothenburg

UNIVERSITY OF GOTHENBURG

Gothenburg 2024

Epidemiological aspects of health and work ability in the Swedish workforce

Adnan.Noor.Baloch@gu.se

ISBN 978-91-8069-671–5 (PRINT)
ISBN 978-91-8069-672–2 (PDF)
https://hdl.handle.net/2077/79340

Printed in Borås, Sweden 2024
Printed by Stema Specialtryck AB

SVANENMÄRKET
Trycksak
3041 0234

To my parents

Mehr-un-Nisa and Noor Muhammad Baloch

Epidemiological aspects of health and work ability in the Swedish workforce

Adnan Noor Baloch

School of Public Health and Community Medicine, Institute of Medicine

Sahlgrenska Academy, University of Gothenburg

Gothenburg, Sweden

ABSTRACT

Background: Many working-age individuals suffer from health issues, impacting their work ability and potentially leading to early workforce exit. More understanding of individual and occupational factors influencing work ability and the use of machine learning techniques in occupational health research is needed.

Aim: This thesis aims to describe and investigate the effect of health and work environment factors on work ability. In addition, it seeks to compare different statistical and machine learning methods for predicting work ability among workers by using demographic and work environment data.

Methods: Study I and II used data from 247 working-age gynecological cancer survivors to study *if* and *how* gastrointestinal syndromes influence the likelihood of disability pension. Study I utilized log-binomial regression, and Study II used counterfactual-based mediation analysis methods. Study III used multiple mediation analysis to investigate stress's long-term impact on work ability among 1432 young adults. Study IV compared the predictive performance of logistic regression with machine learning techniques among 6302 workers with musculoskeletal symptoms.

Results: In Study I, an increased risk of disability pension was found for survivors with radiation-induced gastrointestinal syndrome. However, Study II showed that this increased risk is mediated via different aspects of self-assessed quality of life, including global physical health, physical strength, psychological well-being, and satisfaction with sleep. Study III indicated that the association between perceived stress (at baseline) and work ability among young adults was mediated by stress five years later, feelings of control over one's personal life, work demands affecting personal life, and feeling well-rested upon waking. In Study IV, all machine learning techniques exhibited strong performance, with Extreme Gradient Boosting, AdaBoost, and Gradient Boosting Machines performing best. The calibration varied across techniques.

Conclusion: There exists a complex relationship between health and work ability. Radiation-induced gastrointestinal syndrome influences the self-reported quality of life and disability pension, where quality of life mediates part of the gastrointestinal syndrome's effect on disability pension. Occupational factors mediate the impact of perceived stress on work ability in young adults. Also, feelings of control over personal life are important for maintaining work ability in young adults with stress complaints. The performance of multivariable logistic regression, which utilizes clinically relevant predictors, is comparable to that of sophisticated machine learning algorithms when it comes to predicting disability pensions.

Keywords: Work ability, female cancer survivors, young adults, epidemiology, disability pension, work environment, musculoskeletal symptoms, stress, machine learning.

ISBN 978-91-8069-671–5 (PRINT)

ISBN 978-91-8069-672–2 (PDF)

SAMMANFATTNING PÅ SVENSKA

Bakgrund: Många individer i arbetsför ålder lider av hälsoproblem, vilket försämrar deras arbetsförmåga och kan leda till ett tidigt utträde ur arbetslivet. Det behövs mer kunskap om hur individuella och yrkesmässiga faktorer kan påverka arbetsförmåga samt användningen av maskininlärningsmetoder inom arbetshälsoforskning.

Syfte: Denna avhandling syftar till att beskriva och undersöka effekten av hälso- och arbetsmiljöfaktorer på olika aspekter av arbetsförmågan och jämföra olika statistiska algoritmer och metoder för att förutsäga arbetsförmåga.

Metod: Studie I och II använde data från 247 överlevare av gynekologisk cancer i arbetsför ålder för att studera *om* och *hur* strålningsinducerade gastrointestinala syndrom påverkar sannolikheten för förtidspension. Studie I använde log-binomial regression och studie II använde ramverket för potentiella utfall för medieringsanalyser. I studie III undersöktes stressens långvariga inverkan på arbetsförmågan hos 1432 unga vuxna med hjälp av multipel medieringsanalys. Studie IV jämförde den prediktiva prestandan hos logistisk regression med maskininlärningsmetoder bland 6302 arbetstagare med muskuloskeletala symtom.

Resultat: I studie I fanns en ökad risk för förtidspension hos gynekologiska canceröverlevare med strålningsinducerade gastrointestinala syndrom. I studie II pekade resultaten på att denna ökade risk medieras via olika aspekter av självskattad livskvalitet, inklusive global fysisk hälsa, fysisk styrka, psykologiskt välbefinnande, och tillfredsställelse med sömn. I studie III visade resultaten att sambandet mellan upplevd stress (baslinje) och arbetsförmåga fem år senare hos unga vuxna, medierades av stress fem år senare, arbetskrav som påverkar privatlivet, känslor av kontroll över privatlivet och känslan av att vara utvilad när man vaknar. I studie IV uppvisade alla metoder hög prediktiva förmåga, där Extreme Gradient Boosting, AdaBoost och Gradient Boosting Machines presterade bäst. Kalibreringen varierade mellan olika metoder.

Slutsats: Det finns ett komplext samspel mellan hälsa och arbetsförmåga. Strålningsinducerat gastrointestinalt syndrom påverkar självrapporterad livskvalitet och framtida förtidspension där livskvalitet medierar en del av syndromets effekt på förtidspension. Faktorer i arbetsmiljön medierar stressens inverkan på arbetsförmågan hos unga vuxna. Även känslan av kontroll över privatlivet är viktig för att bibehålla arbetsförmågan hos unga vuxna som

upplever stress. Logistisk regression, med kliniskt relevanta prediktorer, är jämförbar med avancerade maskininlärningsalgoritmer när det gäller att förutse förtidspension.

LIST OF PAPERS

This thesis is based on the following studies, which are denoted in the text by Roman numerals.

I. **Noor Baloch, A.**, Hagberg, M., Thomée, S. Steineck, G, Sandén, H. Disability pension among gynaecological cancer survivors with or without radiation-induced survivorship syndromes. Journal of Cancer Survivorship 16, 834–843 (2022). https://doi.org/10.1007/s11764-021-01077-9

II. **Noor Baloch, A**, Hagberg, M, Thomée, S, Steineck, G, Sandén, H. The physical and psychological aspects of quality of life mediates the effect of radiation-induced urgency syndrome on disability pension in gynecological cancer survivors. Cancer Medicine. 2023; 12: 17377-17388. https://doi.org/10.1002/cam4.6356

III. van Schaaijk A, **Noor Baloch A**, Thomée S, Frings-Dresen M, Hagberg M, Nieuwenhuijsen K. Mediating Factors for the Relationship between Stress and Work Ability over Time in Young Adults. Int. J. Environ. Res. Public Health. 2020, 17(7),2530; https://doi.org/10.3390/ijerph1707253

IV. **Noor Baloch, A**, Sandén, H, Hagberg, M, Adiels, M. Comparable performance of logistic regression and machine learning techniques in predicting disability pension among workers with musculoskeletal symptom. Manuscript.

CONTENTS

ABBREVIATIONS

AI	Artificial intelligence
AUC	The area under the curve
CI	Confidence intervals
EU-OSHA	European Agency for Safety and Health at Work
LISA	Longitudinal Integrated Database for Health Insurance and Labour Market Studies
ML	Machine learning
NDE	Natural direct effect
NIE	Natural indirect effect
PM	Proportion mediated
ROC	Receiver operating characteristic curve
RD	Risk difference
RR	Relative risk (also known as risk ratio)
SCB	Statistiska centralbyrån (Statistics Sweden)
TE	Total effect
WAI	Work Ability Index
WAS	Work Ability Score

1 INTRODUCTION

Work ability is a key factor in occupational functioning and plays a role in determining sustainable employability [1-3]. It holds significance at personal, societal, and political levels [4]. The potential for an enhanced and sustained work life is linked to individuals' work ability. Regrettably, common diseases such as cardiovascular, musculoskeletal, and mental disorders prevalent among the workforce, have an impact on their work ability, thereby reducing their chances of working until the standard pension age [5-8]. Additionally, this issue is further compounded by the aging population and the extension of the working lifespan [9].

Radiation therapy, especially when used to treat gynecological cancer, gives rise to long-lasting conditions that have a substantial impact on gastrointestinal health [10-14], work ability [15, 16], and overall quality of life [11-13, 17]. These gastrointestinal syndromes caused by radiation not only affect physical health but also have consequences for social and occupational functioning [18], possibly leading to the need for disability pensions.

Perceived stress is a prevalent concern in the field of public health and has an impact on work ability [19-22]. The 2022 Occupational Safety & Health Survey by the European Agency for Safety and Health at Work (EU-OSHA) revealed that 27% of workers experience stress, anxiety, or depression due to or worsened by their work [23]. Young workers, typically those below the age of 35 [24], exhibit the highest incidence of stress complaints. High levels of stress can lead to a heightened risk of occupational injuries [25] and, according to the life course perspective, can precipitate burnout, depression, and adverse employment outcomes later in life [26, 27].

Swedish research has contributed to investigating factors that contribute to disease risk [28] and the cost of illness [29] by linking data with nationwide registers. This linkage of data provides a unique opportunity to explore observational and administrative data. However, as a result, complex datasets [30] are generated, which require advanced data analysis methods to extract meaningful insights. Machine learning techniques offer greater flexibility and adaptability in modeling such complex data structures. It is crucial to compare the predictive performance of different machine learning techniques to develop more accurate, robust, and generalizable predictive models in the field of occupational health research. This highlights the importance of using statistical techniques that are better suited for analyzing complicated data structures when examining factors that contribute to decreased work ability among individuals with poor health.

In summary, a significant number of working-age adults face health issues that hinder their ability to work. Therefore, it is essential to understand the pathways through which health problems impact work ability. Unfortunately, there is a lack of high-quality, methodologically rigorous prospective studies that examine the relationship between health and work ability, particularly studies that involve a population-based sample. With regards to data analysis methodology, as far as we know, there is no existing literature that has compared the performance of machine learning techniques to guide their use in predicting work ability.

The objective of this thesis is to examine *if* and *how* radiation-induced gastrointestinal syndromes are associated with disability pension, as well as *how* the association between stress and work ability evolves in young adults. Additionally, it compares logistic regression with six advanced machine learning techniques, namely AdaBoost, Extreme Gradient Boosting, Gradient Boosting Machine, Naïve Bayes classifier, Random Forest, and Support Vector Machine, in predicting disability pensions among workers experiencing musculoskeletal symptoms.

Although the focus of this research is on Sweden, specifically addressing issues related to musculoskeletal symptoms, perceived stress, and cancer, its findings have global relevance.

1.1 WORK AND HEALTH

Work is essential in people’s lives and an important social determinant of health. Participation in working life is associated with better health and is central to social identity and good economic standards [31]. In ideal situations, work offers favorable conditions to promote health and well-being; working enables and improves society's overall health. The national Work-Related Disorder survey from 2020 found that nearly 32% of Sweden's working population, which is approximately 1.6 million out of a total workforce of about 5 million, had reported health issues in the past year that were related to their work [32].

1.2 WORK ABILITY – DEFINITIONS AND MODELS

The precise definition of work ability remains elusive, with existing definitions and operationalizations often being fragmented or imprecise [33]. Generally, work ability refers to the balance between job demands and an individual’s capacity to perform tasks effectively [34]. Understanding the various aspects of work ability is essential for promoting extended work life, fostering employment growth, and enhancing well-being. Work ability, a multifaceted concept, involves factors relevant to an individual's resources to meet job demands. These factors may incorporate subjective elements such as values and attitudes, which can be challenging to assess. Work ability is central in health insurance, as it is the basis for determining benefits and compensations in the event of an illness or disability that impairs one's work ability [35]. Thus, defining work ability to establish legal rights to social insurance benefits is essential.

According to Nordenfeldt [36], work ability contains a range of abilities that enable an individual to perform tasks and achieve goals required by their job, within a suitable work environment. Practical work ability depends on both an individual's capabilities and environmental facilitators. Tengland [37] further elaborates on work ability by differentiating between specific and general work ability. Specific work ability pertains to jobs requiring specialized education or training, while general work ability applies to jobs most people can perform following basic training. This classification of work ability has legislative implications for health insurance regulation, underscoring that both internal factors (individual abilities) and external factors (environmental opportunities) jointly influence work ability. Both definitions highlight that work ability is shaped by internal and external factors and is tied to specific tasks and environments, signifying that work ability is a dynamic interplay between individual abilities, environmental opportunities, and contextual job demands.

Ilmarinen's multidimensional model, "Work Ability House” comprising four floors, is central to the theory of work ability [4, 38], as illustrated in Figure

1.1. It considers work ability as an individual resource in relation to work demands. The resources, which include health, mental and physical capacity,

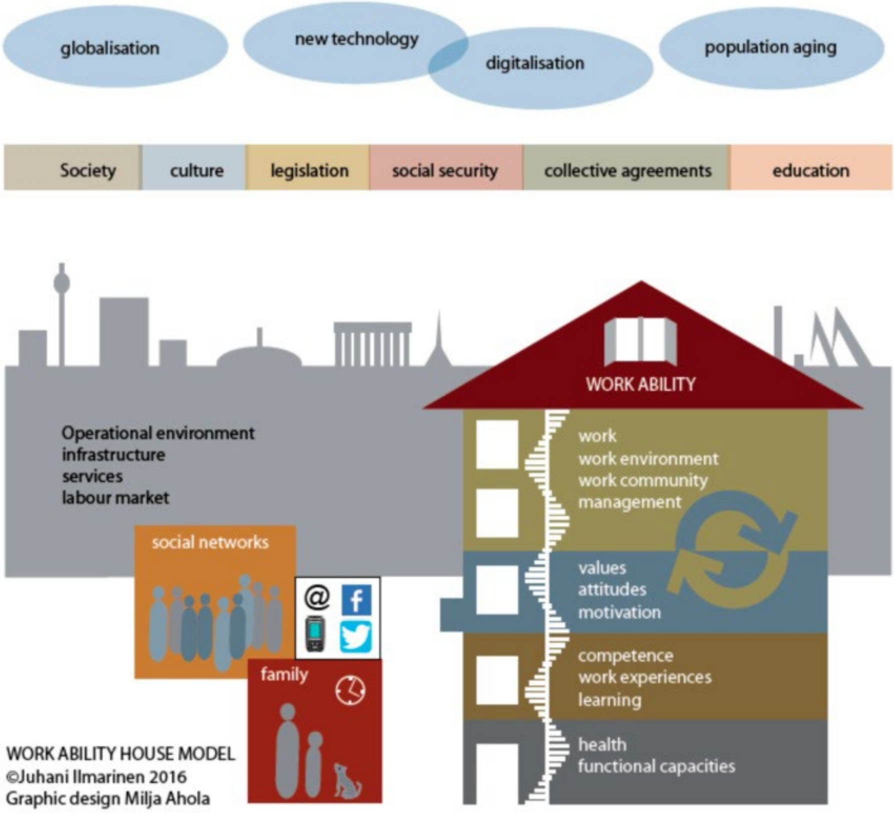

Figure 1.1: The Work Ability House model depicting dimensions of work ability. Original published in: From Work Ability Research to Implementation by Juhani Ilmarinen. https://doi.org/10.3390/ijerph16162882

knowledge and skills, values, and attitudes form the floors of the house. These resources are related to mental, physical, and psychosocial demands at work and the work environment (Figure 1.1). According to Ilmarinen, the Work Ability House exists within a broader context. Work ability extends beyond the workplace and is influenced by factors such as occupational health care and safety, family, and the close social circle (relatives, friends, acquaintances). Furthermore, at the societal level, infrastructure, social, health, and occupational policies, as well as legal, educational, and cultural factors collectively influence work ability and the conditions under which individuals can effectively participate.

The Work Ability Index is widely used to evaluate work ability in occupational health services and scientific studies [39]. It assesses a worker's work demands, health, and resources to meet job demands, providing a holistic image of their work ability. The Work Ability Index and its components are recognized predictors of long-term sick leave, early retirement, disability pension, and unemployment [1-3, 40, 41]. Critics have raised concerns that the Work Ability Index comprises questions related to multiple dimensions such as work ability, disease diagnoses, work impairment, mental resources, and sickness absence. They have concluded that the one-dimensional structure of the WAI needs to be revised for measuring work ability [42]. Additionally, the Work Ability Index includes questions related to people's ability to work, such as their disease diagnoses, injuries, and sickness absence, and their medical diagnosis may overly influence it.

While the WAI is a reliable predictor, large-scale surveys may only use part of the index. Individual WAI items, such as the first item in WAI, namely the Work Ability Score that compares the current work ability to the highest work ability, are strongly associated with WAI and can predict long-term sick leave and disability pension [41, 43, 44].

1.3 THE SOCIAL INSURANCE AGENCY

In a welfare state, the government’s role is to protect the well-being of all citizens, ensuring that they have access to the resources and support they need to lead dignified lives. The Social insurance system in Sweden is a cornerstone of the Swedish social security system and aims to provide financial security during different stages of life. It includes sickness insurance, disability pensions, and parental insurance and benefits. All three categories apply uniformly to individuals habitually residing or working in Sweden. Swedish citizenship is not a prerequisite for eligibility within the social insurance system. The system is financed through taxation and earnings-related Employers’ contributions.

The Social Insurance Agency (Försäkringskassan) is a government agency, which follows state rules and regulations and administers the social insurance system in Sweden. It is responsible for evaluating work capacity and determining eligibility for receiving social insurance benefits and compensation. Since 2007, several changes have been implemented in the

Swedish social insurance system. One of the most noteworthy revisions involved the introduction of the work ability concept, which means that a person's ability to work is assessed in relation to the occupation from which they have been absent. According to statutory regulations, individuals covered by the insurance are eligible for benefits if their work capacity is reduced by a minimum of twenty-five percent due to sickness or disability.

The evaluation of work ability does not take into account factors such as the labor market (e.g. job availability), the individual's financial situation, social support network, or similar conditions. The assessment of work ability is based on medical documentation provided by physicians and is linked to the International Classification of Functioning, Disability, and Health (ICF). A physician establishes a diagnosis and assesses if any functions are impaired and if the impairment has led to a decreased work ability (for performing the current work). This evaluation considers the individual's ability to perform tasks: 1) within the scope of their current employment, 2) with adapted work tasks within the same employer, and 3) within the broader labor market context, depending on the duration of the individual's sick leave.

1.4 DISABILITY PENSION

Disability pension is a register-based commonly used indicator of work-related health. Certain demographic groups, including women and older workers, are more likely to be awarded a disability pension [45]. Factors such as declining health with age and prolonged exposure to workplace hazards contribute to the risk of disability pension [8]. Occupational disparities may also influence the likelihood of disability pension, blue-collar workers are reported to have a higher prevalence of early-age pension compared to their white-collar counterparts [46].

Until 2002 early retirement (referred to as "förtidspension" in Swedish) was offered for individuals between the ages of 16 and 64 years who had a long-term or permanent medically based at least 25% reduction in their work ability [47]. During this period, temporary reduction in work ability was compensated through a sickness allowance (known as "sjukbidrag" in Swedish). The early retirement and sickness allowance was part of the old-age pension system. In 2003, early retirement and sickness allowance became part of the social insurance system and changed their name to Activity compensation (known as

"Aktivitetsersättning” in Swedish) and Sickness compensation (referred to as “Sjukersättning” in Swedish) [48, 49].

Activity compensation is temporary (1 to 3 years) and is only granted to 19-29-year-old individuals whose work ability is reduced by at least one-fourth, as a result not able to work for at least one year in any job in the whole labor market [49]. On the other hand, sickness compensation is permanent, intended for 19-64 years old individuals with a permanent reduction in work ability due to chronic sickness, injury, or a disability [48, 49]. Throughout this thesis, the term disability pension will be used for both types of compensation. Disability pension can be granted in full or partially depending on the proportion of reduction in work ability. To be eligible for disability pension benefits, individuals must meet specific criteria related to their health condition and work ability. A comprehensive medical evaluation is conducted to assess the severity and long-term nature of the sickness/disability. Moreover, the individual's ability to engage in gainful employment is considered, considering factors such as education, skills, and work experience.

1.5 FACTORS AFFECTING WORK ABILITY

In this thesis, work ability serves as the overarching concept encompassing both disability pension and work ability score (WAS). Extensive literature reviews [50, 51] recognize pathways between health and work ability and propose a model for research about work in cancer survivorship, see Figure 1.2 These models were the inspiration for investigations in Studies I and II.

1.5.1 RADIATION-INDUCED GASTROINTESTINAL SYNDROMES

Radiation therapy is an essential treatment option for cancer patients, employed in every other cancer treatment, and in some cases, representing the most effective option for specific diagnoses. Despite the precision of modern radiation treatments, healthy tissues in or adjacent to the radiation field can still be affected. The impact poses tremendous challenges for survivors, particularly when treating cancer in the pelvis, as it can affect vital pelvic organs and functions.

Individuals who have received radiation therapy for pelvic cancers, e.g.,

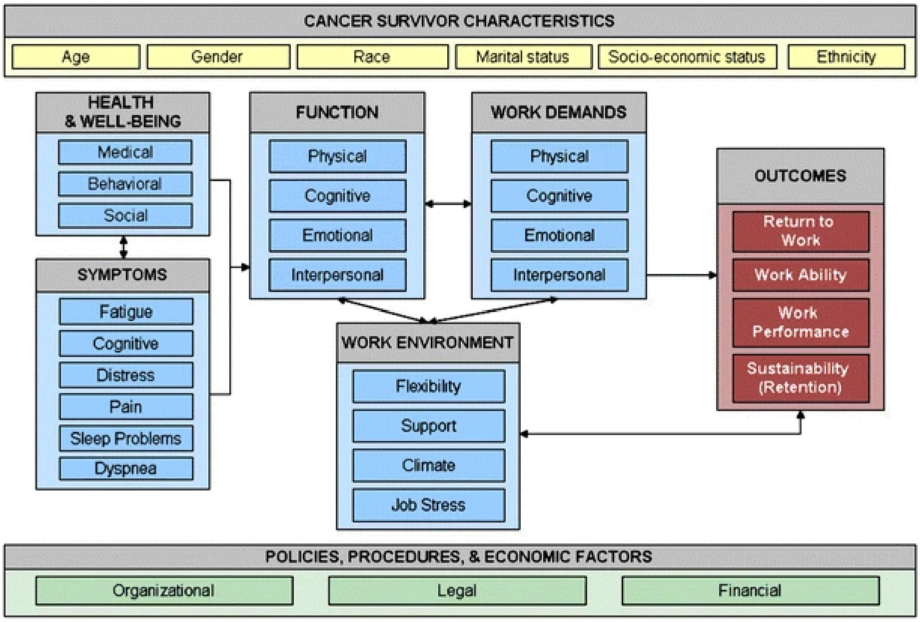

Figure 1.2: The Cancer and Work model. Figure originally published in Feuerstein et al. (2010). Work in cancer survivors: a model for practice and research https://doi.org/10.1007/s11764-010-0154-6. The image is licensed under terms and conditions provided by Springer Nature and Copyright Clearance Center.

gynecological (endometrial, cervical, vulvar, and vaginal), prostate, testicular, rectal, bladder, and anal cancers, are expected to experience varying degrees of dysfunction in the intestinal- and urinary tracts, as well as effects on genital and sexual health many years after treatment. This can lead to lifelong radiation-induced symptoms that cause a decrease in bowel health [10-14] and urinary and sexual health of survivors [52], which may be permanent. Additionally, these symptoms adversely impact the ability to resume regular activities, quality of life [11, 13, 17], and the work ability of cancer survivors [16]. Earlier research has also shown that cancer survivors treated with radiation therapy only or in combination have an increased risk of staying off work than patients who underwent surgery alone [16, 53].

The negative impacts of gynecological cancer and pelvic radiotherapy on work life, social life, and quality of life have been previously documented. Nevertheless, *if* and *how* the radiation-induced gastrointestinal syndromes influence the likelihood of being granted a disability pension remains unclear.

The impact of radiation-induced gastrointestinal syndrome on survivors' transition to disability pension is not fully understood, warranting exploration of mediating pathways. Identifying how these syndromes influence disability pension risk can reveal direct and indirect consequences, thereby informing interventions for cancer survivors' rehabilitation. Enhancing our understanding of potentially modifiable mediating factors may aid in preventing early retirements due to disability in a patient group that has long been overlooked. It is valuable to explore every opportunity for improvement to avoid unfavorable outcomes associated with radiation-induced gastrointestinal syndrome.

1.5.2 PERCEIVED STRESS

Stress can mean different things to different individuals [54]. It can refer to stress exposure or stress reactions [54], hence leading to confusion. It is a multifaceted concept that encompasses several dimensions. To fully grasp it, we must first understand the different objects that constitute it. Stress stimuli include external exposures; stress experience represents the individual perception of the stimuli; general stress response is the body's physiological response and perceived stress is the feedback from the stress response [55].

Perceived stress is an individual's subjective appraisal of the stressfulness of a particular situation. The sum of this perceived stress is measured frequently in questionnaires in occupational health research. It is an essential element of many anxiety scales and questionnaires on health [55]. It can be influenced by many factors, including personality traits, coping strategies, social support networks, and work environment factors. Moreover, the subjective nature of perceived stress underscores its variability across individuals and contexts, as what may be perceived as stressful for one person may not produce the same response in another [55]. The consequences include physical and mental health issues [27].

The concept of perceived occupational stress encompasses the reactions individuals manifest when confronted with work-related tasks and pressures that exceed their cognitive and functional capacities, thereby impeding their ability to work optimally. Understanding perceived stress influence is crucial as it plays a pivotal role in psychological well-being, physical health outcomes, work ability, and overall quality of life [19-22].

Thus, exploring the intricate pathways of perceived stress that influence work ability provides valuable insights into the mechanisms and informs targeted interventions to promote work ability.

1.5.3 WORK ENVIRONMENT

Employees who have physically demanding jobs are at a higher risk of suffering from impaired work ability, sickness, and early retirement from work [1, 5, 56, 57]. Strenuous work environment such as physical workload and ergonomic factors such as heavy lifting, working with hands lifted to shoulder height or higher, and working with the twisted or bent forward are associated with work ability [56] and disability pension [1, 58-61]. According to the European Working Conditions Telephone Survey 2021 (EWCTS), half of the workers at least sometimes worked in tiring or painful positions, whereas one in three workers carried heavy loads [62].

Psychosocial work environment refers to social and organizational factors that can influence the physical or mental health of workers. Poor psychosocial work environments are strongly associated with work ability [1, 56] and disability pension [1, 60, 61]. A poor work environment also has economic and societal impacts, such as increased health expenses and lower productivity.

The Swedish Work Environment Authority (in Swedish, Arbetsmiljöverket) is a government organization that ensures workplaces in Sweden follow work environment regulations. Statistics Sweden (in Swedish, Statistiska centralbyrån) on behalf of The Swedish Work Environment Authority conducts the biennial Swedish Work Environment survey [63]. The survey, which supplements the regular Labour Force Survey by Statistics Sweden, is conducted via telephonic interviews and postal questionnaires. Following the Labour Force Survey, employed (paid job, self-employed or family business) individuals aged between 16 to 64 years, and not on a long-term sickness absence were invited to participate in the Work Environment Survey. The survey is conducted during October - December [63].

1.5.4 SELF-ADMINISTERED QUESTIONNAIRE IN EPIDEMIOLOGICAL STUDIES

The self-administered questionnaire is the most frequently used method to measure health, work environment, and work-related disorders in epidemiological studies and national surveys. In reaching a large and diverse population, they are easier to administer and more cost-effective than interview methods [64]. They also allow individuals to share their subjective perceptions, experiences, and attitudes covering various dimensions, including physical, psychosocial factors, and organizational, providing a comprehensive view. However, they also pose some limitations, such as response bias and lack of objectivity. The preceding literature review revealed a limited to moderate level of agreement regarding *self-reported illness* and *expert assessment based on clinical examination* [65]. Studies using *symptom self-reporting* to predict *expert assessment* often found moderate-to-high sensitivity but moderate-to-low specificity. Single-question studies typically showed high specificity but low sensitivity [65]. Questionnaires may also offer an incomplete picture and only reach some workforce segments. Furthermore, reliance on self-disclosure may hinder the disclosure of sensitive information. Despite these drawbacks, self-administered questionnaires remain practical in large surveys.

1.6 PREDICTION MODELS IN OCCUPATIONAL HEALTH

In a broader context, prediction models serve as valuable tools for clinical practice and research [66]. In occupational health, these models may play a vital role in directing preventive interventions toward workers at heightened risk of having or developing a particular occupational disorder or disease. In clinical settings, prediction models offer patients and their healthcare providers insights regarding the likelihood of a prognostic or a diagnostic outcome. Prognostic assessments, for instance, may aid in planning an individual's remaining work life in diminished work ability with an early exit from work life or instill optimism for a return to work.

Prediction models are firmly established within clinical medicine [67] but are comparatively new in occupational health. They produce an individualized likelihood of a forthcoming event, such as disability pension (predictive model), or the existence of a condition (diagnostic model). Prediction models try to reply to inquiries like if having a high amount of physical work and a male gender leads to a higher risk of receiving disability pension or what the risk of receiving disability pension is under such circumstances. Predictive models gather data from various sources, including individual characteristics, work-related factors, and disease-related variables, to generate forecasts. They generally consider a moderate number of predictors, ranging from 2 to 20 [66].

Occupational health research often involves complex datasets characterized by nonlinear relationships, high-dimensional features, and variable interactions. Traditional statistical methods like logistic regression may not have the capability to capture the intricate patterns in such datasets. Machine learning techniques, including ensemble methods that combine the predictions of multiple individual models to enhance overall performance, may offer more flexibility and adaptability in modeling complex data structures [68].

1.7 STATISTICAL MODELS FOR PREDICTION

Both classical statistical methods and modern machine learning (ML) techniques are utilized to develop prediction models [69]. Machine learning methods are gaining attention across clinical practice, epidemiology, and medical research, facilitating the analysis of various types of health data [68].

Maximum likelihood logistic regression has been a reliable and interpretable tool for binary classification problems, providing an easily interpretable and transparent result. Furthermore, it is computationally less intensive than machine learning methods, rendering it more practical for larger datasets. However, it may need to be improved in situations with complex, high-dimensional data, or non-linear relationships.

Machine learning models are adept at analyzing numerous potential predictors by dynamically identifying complex patterns, which may encompass high-order interactions and nonlinear relationships. They also offer expanded possibilities for variable selection and prediction modeling, with the choice of technique significantly impacting prediction models' accuracy, interpretability, and efficiency. An advantage of machine learning techniques, exemplified by the Gradient Boosting Machine, is leveraging the sequential correction of errors made by previous models, enhancing their effectiveness in capturing complex patterns. Despite being useful in prediction modeling, some ML techniques are criticized for their lack of interpretability [69].

The utilization of machine learning in occupational health research remains relatively limited [70] and comparative studies on its performance are scarce. Moreover, a knowledge gap exists regarding the potential application of machine learning in forecasting work-life outcomes such as disability pensions. When using machine learning in research on occupational health, it is crucial to understand the strengths and weaknesses of different machine learning methods to make informed decisions regarding model selection and interpretation.

Interpretability is another critical consideration in occupational health research, primarily when predictive models inform workplace interventions and policy decisions. Logistic regression models offer straightforward interpretability, allowing researchers to identify predictors' relative importance and assess associations' direction and strength with outcomes. In contrast, some machine learning techniques, may be opaque and exhibit a low degree of explain ability, making them challenging to interpret due to their complex architectures and black-box nature.

Researchers can assess the trade-offs between bias and variance of different machine learning techniques. In addition, they can also balance the need for predictive accuracy with the importance of transparency and explainability in occupational health modeling. Comparative evaluations of machine learning techniques, including logistic regression, contribute to the reproducibility and external validation of predictive models in occupational health research.

By transparently reporting the detailed description and explanation of model building, model architecture, training, and testing procedure, and algorithm performance metrics, researchers enable other investigators to replicate their findings and assess the generalizability of predictive models across different populations, settings, and contexts. This promotes scientific rigor and reliability, enhancing the credibility and utility of predictive modeling in occupational health.

At the same time, researchers can identify the most effective models for predicting occupational health outcomes such as work-related disorders, or disability pension. By leveraging the strengths of diverse algorithms and methodologies, researchers can advance our understanding of occupational health outcomes, inform evidence-based interventions, and ultimately improve the health and well-being of workers worldwide.

2 AIM

This thesis aims to describe and investigate the effect of health and work environment factors on work ability. In addition, it compares different statistical algorithms and methods for predicting work ability by using demographic and work environment data.

The specific aim of each study is:

1. To determine the extent to which the intensity of different radiation-induced gastrointestinal syndromes, each decreasing intestinal health, affects the likelihood of being granted a disability pension.
2. To investigate and quantify the mediating role of physical and psychological aspects of quality of life on radiation-induced urgency syndrome and future disability pension association.
3. To investigate how the relationship between stress and work ability is mediated in young adults over time by recovery, the influence of job demands on private life, feelings of control over private life, and physical activity during leisure time.
4. To compare the relative performance of logistic regression and six modern machine learning algorithms in predicting disability pension among workers with musculoskeletal symptoms in the Swedish workforce.

3 MATERIALS AND METHODS

The aim of this thesis was explored with four observational prospective longitudinal cohort studies, utilizing quantitative research methods. They were based on three different data sources with a focus on factors that hinder or promote work ability. The study object was individual, with an individual health perspective utilizing data on an individual level. The research was in the discipline of public health. In Studies I, II and IV, the outcome was an all-cause disability pension including all diagnoses. Both self-administered questionnaires and register data were used in the analysis. The specific details of the study designs, study populations, procedures, and variables used are elaborated upon for each study. A comprehensive overview is presented in Table 3.1

3.1.1 STUDY I AND II

Studies I and II were clinical observational prospective cohort studies investigating *if* and *how* radiation-induced gastrointestinal syndromes are associated with disability pensions. These studies use patient-reported information obtained via questionnaires and disability pension data from the official register, the Longitudinal Integrated Database for Health Insurance and Labour Market Studies (LISA) [47].

3.1.1.1 THE GYNECOLOGICAL CANCER COHORT

A clinical patient group comprising 1,800 women who underwent external pelvic radiotherapy for gynecological cancer between 1991 and 2003 was identified using medical records. The radiotherapy was administered at

Table 3.1:Overview of materials and methods of the Studies I-IV

Study	Cohort	Study population	Sample (n)	Baseline	Follow-up	Outcome	Analysis
I	GynCancer	Women treated with pelvic radiotherapy	247	2006	2008	Binary disability pension	Logistic regression
II	GynCancer	as above	247	2006	2008	as above	Causal mediation analysis
III	WAYA	Young adults[a]	1432	2012	2017	Self-reported work ability	Linear regression-based mediation analysis
IV	Work environment survey	Employed persons in Sweden[b]	6302	2011 /2013	2012 /2014	Binary disability pension	Machine learning techniques

[a]Young adults were 25-29 years old in 2012
[b]Employed persons, aged 16-64 years

Jubileumskliniken, Sahlgrenska University Hospital in Gothenburg (Sweden), or Radiumhemmet, Karolinska University Hospital, Stockholm (Sweden). Out of this group, 1303 were alive in 2004. Of these, 823 cancer survivors met the inclusion criteria, i.e., to have been born in 1927 or later, be proficient in reading and understanding Swedish, and have no recurrent disease.

From January to October 2006, these 823 cancer survivors were contacted via an introductory postal letter detailing the study's objectives. Subsequently, each survivor was called by telephone. Among those who provided informed consent during the telephone conversation (n=723) were sent a postal questionnaire. A thank-you card and a reminder were also sent approximately three weeks following the questionnaire's mailing. In cases where it was deemed appropriate, a reminder telephone call was made. In total, 650 cancer survivors returned the fully completed questionnaire. However, seven individuals were excluded due to missing information regarding gastrointestinal health, and 20 were excluded due to having a bowel stoma. The recruitment process is outlined in detail in the flowchart provided in Appendix 1.

The postal questionnaire encompassed 351 questions covering various topics, including demographics, concurrent illnesses, comorbidities, psychological issues, quality of life, and sexual function. Respondents were also asked about symptoms related to the abdomen, gastrointestinal system, genitals, legs, pelvic bones, and urinary bladder. Moreover, the questionnaire included inquiries about these symptoms' presence, severity, and duration. Earlier studies have comprehensively defined the enrollment and data collection procedure [71-73]. In addition, the process of developing and validating the questionnaire has also been thoroughly documented [73].

3.1.1.2 THE LONGITUDINAL INTEGRATED DATABASE FOR HEALTH INSURANCE AND LABOUR MARKET STUDIES (LISA)

All adult Swedish residents aged ≥15 years (before 2010 aged ≥16 years) who are listed in the Total Population Register as of December 31st each year and are covered by the social insurance system, are encompassed by the Swedish Longitudinal Integrated Database for Health Insurance and Labour Market Studies (LISA, Longitudinell Integrationsdatabas för Sjukförsäkrings- och

Arbetsmarknadsstudier) [47]. The Swedish Longitudinal Integrated Database for Health Insurance and Labour Market Studies combines information on individuals by retrieving data from official Swedish registries. The unique personal identity number possessed by every individual listed in the Total Population Register facilitates the data linkages.

3.1.2 STUDY III

Study III was an observational longitudinal cohort study investigating the pathways through which stress can affect work ability over time among young adults, using a questionnaire administered in 2012 and 2017.

3.1.2.1 THE WAYA COHORT

The Work Ability Young Adults (WAYA) cohort was established in 2007 to understand better the work environment and lifestyle factors that influence *well-being* and sustainable *physical and mental* work ability among young adults [74, 75]. The researchers invited 20,000 individuals aged 20-24 randomly selected from the total population register, with equal representation of men and women, to participate. Half of the cohort resides in Västra Götaland county, Sweden, and the other half in the rest of Sweden. The age range of 20 to 24 years used in WAYA aligns with health research which defines young adults as people between 16 and 25 years old [76]. Of the selected individuals, 7,125 (2,778 men and 4,347 women) responded to the initial comprehensive postal questionnaire. The questionnaire focused on questions regarding health, work ability, information and communications technology use (mobile phone and computer use), occupational and recreational exposures, demographic details, general life situation, and physical and psychosocial work demands. To encourage participation, a lottery ticket worth 10 SEK (ca 1 euro) was enclosed with the cover letter, which could be used irrespective of survey completion.

Two reminder letters were subsequently mailed to the participants. Follow-up surveys were conducted in 2008 and 2012 on the subset of respondents who had expressed their willingness to participate in future studies during the baseline survey in 2007 (5734 individuals). They were invited to complete an identical self-administered online questionnaire. The data collection process closely mirrored that of the baseline survey, but with an additional third reminder that offered the choice to receive a paper version of the questionnaire and two cinema tickets as an extra incentive.

The response rate notably increased to 73% in 2008, with a total of 4,163 participants, consisting of 2,705 women and 1,458 men. Additionally, the response rate reached 55% in 2012, with 1,739 women and 999 male participants. The 10-year follow-up survey conducted in 2017 was sent to the baseline survey respondents from 2007, with 1,814 women and 1,119 men responding (a response rate of 41%). Participants had the option to respond online for all surveys if they preferred. These follow-up surveys gathered additional information on various aspects of participants' lives, such as mobile and computer usage, work demands and exposures, general life situation, health, work ability, and physical activity. A more detailed description of the WAYA can be found in previous publications [77-79]. An illustrative flowchart is available for reference in Appendix 2.

3.1.3 STUDY IV

In Study IV, we compared the *discriminative* performance and *reliability* of logistic regression's risk predictions *(calibration)* with six modern machine learning algorithms for predicting binary disability pensions. This study uses work environment data from the National Work Environment Survey and disability pension data from the official register, the Longitudinal Integrated Database for Health Insurance and Labour Market Studies (LISA) (see 3.1.1.2 LISA).

THE WORK ENVIRONMENT SURVEY

The Swedish Work Environment survey is bi-yearly and supplementary to the Labour Force Survey (LFS) [80], conducted by Statistics Sweden (SCB) on behalf of the Swedish Work Environment Authority (Arbetsmiljöverket) [80]. The Swedish Labour Force Survey is comprehensive and nationally representative. The Labour Force Survey is designed to obtain information on the labor market and related issues through structured questionnaire-based telephone interviews with a representative sample of the Swedish workforce (15–74 years of age). The data obtained from the Labour Force Survey is included in the official statistics and is an essential source of information for government agencies, policymakers, and researchers interested in understanding the composition and characteristics of the Swedish labor force.

LFS follows the International Labour Organization (ILO) conventions, recommendations, and the European Union's guidelines. Respondents of LFS falling within the age range of 16 to 64 years, who were gainfully employed (including self-employed) and not currently on sickness absence or maternity leave, were asked an additional 25 questions concerning their work environment [63]. Subsequently, they also receive a postal or an online questionnaire with additional questions emphasizing the health and safety aspects of the work environment. For example, workplace hazards, demands and influence, employee exposure to various risks, work environment management, occupational health, and safety practices, physical or psychosocial and mental problems that the work may have caused, and the overall work climate. The Work Environment survey highlights how workers perceive their working environment and the potential *(physical and mental)* problems it causes. The data from the Work Environment survey are used as a basis for the Swedish Work Environment Authority's activities and by decision-makers at various levels within organizations, government authorities, researchers, and the government and parliament. In 2011 and 2013, the survey received responses from 7926 and 4774 individuals (telephone interview and postal questionnaire), among the 15828 and 9810 invited individuals, respectively.

DATA PROCESSING

Approximately half of the workers reported experiencing musculoskeletal symptoms, which were described as pain in one or more of the following body parts: (1) upper back or neck, (2) lower back, (3) shoulders or arms, and (4) wrists or hands. This pain was experienced at least one day a week (1 day out of 5) or more frequently over the past three months. The data from workers with musculoskeletal symptoms, collected in the 2011 and 2013 surveys, were combined. Missing data were imputed with R-package MICE, using predictive mean matching [81] to mitigate any potential negative impact on performance. To address the issue of multicollinearity, variables with a strong correlation (Spearman or Polychoric correlation ≥ 0.70, used when appropriate) were identified. The variable with the highest average correlation with other variables was subsequently eliminated.

In Study IV, the term "model" is used to denote a varying number of predictors. Three distinct models, each with 7, 13, and 32, respectively. The first model incorporates seven variables that have been carefully selected by an occupational health physician and a professor with extensive research and practical experience. The second model expands upon the first by including additional demographic variables. Finally, the third model employs Boruta, a machine learning technique [82], to select a total of thirty-two variables from a pool of forty-eight candidate variables covering various work environments and demographic factors. These candidates were selected from the Work Environment Survey to provide a comprehensive representation of physical and psychosocial working conditions. The English translation of all these questions (predictor variables) is provided in Appendix 3.

To ensure robustness and generalizability, the models underwent both internal validation through 5 repeats of 10-fold cross-validation and external validation through temporal validation [66, 83]. For external validation, the earlier survey data (from 2011) was utilized to develop the models, while the more recent survey data (from 2013) was used to assess their performance.

3.2 VARIABLES USED

3.2.1 EXPOSURES AND MEDIATORS

RADIATION-INDUCED SURVIVORSHIP SYNDROMES

In a previous study, a modified factor analysis method was applied to data self-reported by gynecological cancer survivors and their matched controls. A total of 28 symptoms, indicative of a decline in intestinal health, were discovered to be linked to six distinct factors [72]. These factors were collectively labeled "Radiation-induced survivorship syndromes" by Steineck et al. They included urgency syndrome, leakage syndrome, constipation, excessive gas discharge syndrome, excessive mucus discharge syndrome, and blood discharge syndrome.

A statistically significant difference was observed between survivors and controls regarding factor score quantiles across all factors, except for constipation. Cancer survivors were classified as having a specific syndrome if their factor loading (syndrome intensity) for that syndrome exceeded the 95th percentile of the controls. It's worth noting that cancer survivors could be categorized as having multiple syndromes. The gastrointestinal symptoms involved in the syndromes are detailed in Appendix 1.

QUALITY OF LIFE

Quality of life (QoL) was assessed through self-reported measures using single-item 7-point visual digital scales. These scales have demonstrated a strong correlation with established instruments [84, 85]. An example question posed to participants was, "*How would you rate your quality of life over the past six months?*" Responses ranged from 1, indicating "*no quality of life*," to 7, signifying "*the best possible quality of life*".

Responses were dichotomized based on the median value into two categories: "*low to moderate*" (1–5) and "*high*" (6–7) for most items. However, items assessing depression and anxiety were dichotomized differently, with "*low to moderate*" defined as scores 1–3 and "*high*" as scores 4–7.

Detailed questions about quality of life and the corresponding options are provided in Appendix 1.

PERCEIVED STRESS

Stress levels were determined through a single-item question from the QPS-Nordic questionnaire. The question was subsequently validated against the Maslach Burnout Inventory and the Mental Health Subscale of the Short-Form-36 (SF-36), which effectively measured stress in various work-life scenarios [86].

An explanatory note defining stress was included with the question: "*Stress means a situation in which a person feels tense, restless, nervous, or anxious, or is unable to sleep at night because his/her mind is troubled all the time. Do you feel this kind of stress these days?"*. Respondents were asked to rate their stress level on a 5-point Likert scale, where a score of one indicated "*not at all*" and a score of five represented "*very much*".

MEDIATORS IN STUDY III

The dimension of recovery was assessed through a question on sleep quality: "*How often in the last 30 days have you woken up feeling well-rested?*". The response options ranged from "*never*" (coded as 1), "*once/a few times per* *month*" (coded as 2), "*several times per week*" (coded as 3), to "*almost every day*" (coded as 4).

The interference between work and home life was measured using a question that examined the negative impact of work/studies on personal life: "*Do the requirements of work/studies affect your personal life (leisure, home, and family life) negatively?*". The response options were "*very rarely*," "*quite rarely*," "*sometimes*," "*quite often*," and "*very often*," which were coded from 1 to 5, respectively.

Control over private life was evaluated with a single item: "*I feel I have control over and can handle things that happen in my private life*." The responses ranged from "*poorly*," "*rather poorly*," "*well*," or "*very well*," and were coded as 1 to 4, correspondingly.

The assessment of physical activity during leisure time was conducted using a single question that inquired about the average level of activity over the past twelve months: "*How physically active are you, and how much do you move in your spare time?*". Response options spanned from sedentary activities (e.g., reading, watching TV) to vigorous physical activity (e.g., running, skiing). This allowed for a thorough evaluation of the levels of physical activity during leisure time.

3.2.2 OUTCOME

DISABILITY PENSION

The disability pension system in Sweden provides financial assistance to individuals whose work capacity is diminished due to health impairments. In studies I, II, and IV, data about disability pensions were procured from the Longitudinal Integration Database for Health Insurance and Labor Market Studies (LISA) [47] , using the unique personal identity number assigned to each study participant.

The eligibility assessment for a disability pension involves evaluation by a physician, who provides a medical statement detailing the illness and its impact on the individual's work capacity. Moreover, homemakers, students, and unemployed individuals are also entitled to disability pensions. Additionally, individuals aged 62 or above have the option to combine their disability pension with an early old-age pension [49]. In this thesis, the outcome variable about the disability pension was binary, indicating whether the pension was "granted" or "not granted."

WORK ABILITY SCORE

The Work Ability Score (WAS) was utilized as a measure of work ability. This score represents the first item of the Work Ability Index. The question posed was: "*Assume that your work ability at its best has a value of 10 points. How many points would you give your current work ability?*" Participants were asked to rate their work ability on a scale from 0, which signifies a complete lack of work ability, to 10, which denotes the highest level of work ability they have ever experienced. WAS has demonstrated satisfactory convergent validity in comparison to the complete Work Ability Index. [43].

3.3 MEDIATION ANALYSIS

Occupational health research transcends the mere determination of the existence of an effect. It also seeks to comprehend the operational mechanisms of these effects. The former pertains to *if* a relationship exists, while the latter refers to *how*, the transmission method of an effect. Consequently, the interest in identifying causal mechanisms has grown among researchers. Mediation analysis, a sophisticated regression technique, has emerged as a powerful tool in this endeavor. It is employed to unravel the intricate pathways linking an exposure (X) to an outcome (Y), Figure 3.1. Simultaneously, it provides an estimation of the relative magnitude of these diverse pathways. The fundamental idea lies in assessing the indirect effect of the independent variable on the outcome variable via the mediator while controlling for potential confounding variables.

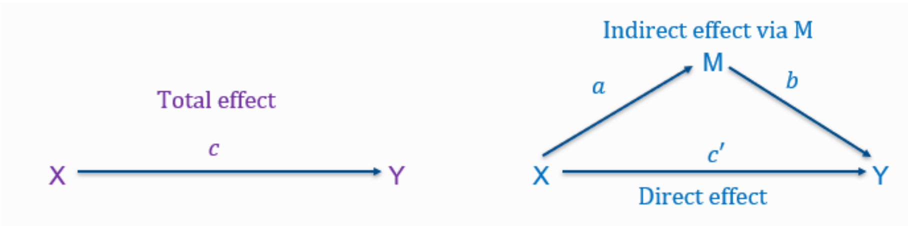

Figure 3.1: A Simple Mediation Analysis Framework, the relationship between variables X (the independent variable), Y (the dependent variable), and M (the mediator variable). **Left**: The total effect of X on Y (c path). **Right:** The total effect 'c' is divided into the direct effects (c´ path) and the indirect effect via M (a and b paths). The indirect effect (a*b), represents the portion of the total effect that is mediated by M.

3.3.1 CLASSICAL MEDIATION ANALYSIS

Classical mediation analysis, popularized by Baron and Kenny [87], is a widely employed regression-based technique aimed at sorting out the underlying mechanisms through which an exposure variable (X) influences an outcome variable (Y). Based on linear regression analysis, Baron and Kenny's seminal framework outlines several sequential steps to establish mediation. These include demonstrating significant associations between the independent variable and outcome (a), mediator and outcome (b), as well as the independent variable and mediator (c). Total mediation is confirmed when the direct effect of the independent variable on the outcome becomes non-significant upon accounting for the mediator.

Traditional mediation analysis offers a structured approach for investigating simple causal pathways, thereby offering insights into the underlying processes that drive observed relationships, but it does have certain limitations [88-90]. Such as, they assume linear relationships between variables, and may not be well-equipped to handle situations involving multiple mediators operating simultaneously or sequentially.

3.3.2 COUNTERFACTUAL-BASED MEDIATION ANALYSIS

Classical methods of mediation analysis [87] may also encounter certain limitations. These limitations arise particularly when handling binary outcomes or mediators. Additionally, where there is confounding between the mediator and the outcome, or when interactions exist between the exposure and the mediator variables [90]. The presence of an interaction suggests that the exposure heightens the vulnerability to the mediator's effect. Incorporating exposure-mediator interactions often augments the magnitude of indirect effects, thereby enhancing the power to detect mediated effects and allowing for greater model flexibility [91]. Consequently, exposure-mediator interactions should generally be incorporated and only excluded if they have a minimal impact on estimates [91, 92].

Counterfactual-based mediation analysis approach addresses these issues [92, 93]. This approach assumes adjustments for confounding variables in the associations between (a) exposure and outcome, (b) mediator and outcome, and (c) exposure and mediator. Moreover, it presumes no direct effect of exposure confounding the mediator-outcome association. Temporal ordering is also assumed [93].

Sensitivity analysis techniques are used to assess the robustness of results in the presence of unmeasured confounding. These techniques enable the evaluation of the combined effect of the confounder-outcome and confounder-mediator associations required to explain indirect or direct effects [94].

3.4 STATISTICAL MODELING

In statistical modeling, two distinct cultures exist based on the foundational principles underlying their approaches [95]. One relies on utilizing stochastic models, assuming that data are generated from known probability distributions. This approach relies on the assumption that the data adhere to specific statistical distributions, such as normal, Poisson, or exponential distributions. In this paradigm, researchers employ classical statistical methods like linear regression to analyze and interpret data. It assumes a linear relationship between variables. While it offers simplicity, linear regression is inadequate in scenarios with non-linear effects.

The second culture assumes algorithmic models, treating the data mechanism as unknown. It embraces machine learning techniques to analyze data without assuming a specific underlying distribution. Machine learning algorithms, such as Random Forests, Support Vector Machines, and Neural Networks, are adept at uncovering complex patterns and relationships within data, even without pre-defined distributional assumptions. Unlike traditional statistical methods, machine learning algorithms learn from the data itself, iteratively refining their models to optimize predictive accuracy or classification performance. This flexibility and adaptability make machine learning particularly well-suited for analyzing large and diverse datasets. In real-world contexts, with non-linear relationships between variables, polynomial regression, and machine learning algorithms, are often employed to study the non-linearity.

Classification typically involves predicting categorical outcomes, while regression pertains to predicting continuous outcomes. Numerous machine learning algorithms initially designed for classification tasks have been modified to address regression problems, and vice versa. This adaptability underscores the versatility of machine learning methodologies across diverse predictive tasks.

Regardless of the modeling approach, the sample size and statistical power are crucial considerations. This aspect becomes particularly important while examining the number of outcome events, especially in relation to binary outcomes. According to existing literature and rules of thumb, a minimum of approximately ten-fifteen events per parameter is necessary for a binary regression method (logistic regression, log-binomial) to ensure robust statistical inference [66, 96]. Machine learning techniques, generally necessitate more than ten times the number of events per variable compared to classical modeling techniques such as logistic regression to attain a stable area under the curve (AUC) [97].

In scenarios where the outcome is exceedingly rare or exceedingly common, a larger sample size becomes imperative to achieve the desired number of events. Both, the overall sample size, and number of events play a pivotal role. Nonetheless, it is vital to acknowledge that the optimal sample size hinges on the specific context and objectives of the study.

3.4.1 LOG-BINOMIAL REGRESSION

The log-binomial model offers a practical method for calculating relative risk and risk difference, for binary outcomes [98, 99]. It predicts the probability (p) of the outcome based on a set of independent (predictor/background) variables, assuming a binomial distribution function and a logarithmic link function, as shown in Equation 1.

$\text{Log}(p) = \sum_{i=0}^{j} \beta_i X_i$ Equation 1: Log (p) is a linear function of regression coefficients.

In the interpretation of the coefficients, the βi values represent differences in log risks (*Log(p)*). Generally, the log-binomial model provides an unbiased estimate of relative risk, though it has some limitations [98].

3.4.2 LOGISTIC REGRESSION

Logistic regression is a well-known statistical method used to model the relationship between a categorical dependent(outcome) variable and one or more independent (predictor/background) variables [100]. Logistic regression is a valuable tool that predicts categorical outcomes based on a set of explanatory variables, particularly useful for binary outcomes. Logistic regression estimates the probability of the outcome variable based on the values of the predictor variables. The distinction between logistic regression and log-binomial regression lies in the link function: logistic regression employs the logit function, while the log-binomial model uses the log function. One immediate implication of this is the difference in interpretation of the coefficients, the βi values denote differences in log odds, as presented in Equation 2.

$\text{Log}\left(\frac{p}{1-p}\right) = \sum_{i=0}^{j} \beta_i X_i$ Equation 2: Logit of probability *(p)* is a linear function of regression coefficients.

The maximum likelihood method fits the model producing a predicted probability (p) of outcome. If the prediction model is good, it produces a predicted probability *(p)* close to zero for individuals without an outcome and a predicted probability *(p)* close to one for all individuals with the outcome. The logistic regression model estimates coefficients for each independent variable, indicating the direction and magnitude of their impact on the log odds of the outcome variable. The exponential of the coefficient (exp (β_i)) provides the odds ratios for the effect of an exposure *Xi*, providing insights into how the likelihood of the outcome changes with unit changes in the predictor variables.

3.5 MACHINE LEARNING TECHNIQUES

Machine learning (ML) encompasses a multidisciplinary domain integrating computer science, statistics, and data science [69, 101, 102]. Various machine learning algorithms have been developed to address the diverse data and problem types encountered in different machine learning scenarios [69]. It enables computers to enhance their performance autonomously through learning from data. This multidisciplinary approach is vital in today's ever-evolving technological landscape with application in many domains of life [101]. In medical practice, machine learning is mainly used for clinical decision support and medical imaging analysis. Clinical decision support tools aid healthcare providers make treatment decisions, prescribe medications, assess health, and address patient needs by offering rapid access to relevant information and research. Additionally, machine learning tools in medical imaging analyze CT scans, x-rays, MRIs, and other images to detect abnormalities or lesions that may be missed by human radiologists.

3.5.1 OPTIMIZING THE PERFORMANCE OF MACHINE LEARNING METHODS

In machine learning, achieving optimal model performance requires a delicate balancing act. While model parameters are obtained directly from the data during the training phase, there exists a distinct set of vital parameters referred to as hyperparameters [103]. These hyperparameters significantly influence the behavior and performance of the technique. Unlike model parameters, hyperparameters are not learned from the data but are predetermined before the training process. The process of hyperparameter tuning aims to identify the optimal values for these configuration variables, which in turn maximizes the model's performance. Tuning involves an iterative search for the most suitable combination of hyperparameters. Grid Search, Random Search, and Bayesian optimization are among the strategies commonly employed for this purpose [103]. Hyperparameter tuning is a blend of art and science. Skillfully adjusting these configuration variables allows one to unlock the model's full potential, leading to improved accuracy, robustness, and generalization.

3.5.2 ENSEMBLE METHODS

An ensemble method in machine learning refers to a technique that combines the predictions of multiple simpler models to enhance overall performance [69]. The rationale behind ensemble methods is rooted in the concept that aggregating diverse models can compensate for the weaknesses of unique models and yield more robust and accurate predictions. There are two main types of ensemble methods: bagging and boosting.

Ensemble methods, through their ability to reduce overfitting, increase model generalization, and improve predictive accuracy, have become popular in various machine learning applications. Standard algorithms that employ ensemble techniques include Random Forest (a bagging method) and AdaBoost, Gradient Boosting, and Extreme Gradient Boosting (XGBoost) (boosting methods). The versatility and effectiveness of ensemble methods make them valuable in the machine learning practitioner's toolkit.

3.5.3 BAGGING (BOOTSTRAP AGGREGATING)

In bagging, multiple instances of the same machine learning algorithm are trained independently on different subsets of the training data [104]. Training data refers to the dataset used to teach a machine learning model, enabling it to learn patterns and make accurate predictions. Each subset is created by random sampling with replacement from the original training dataset, known as bootstrap sampling. The predictions of each model are then combined, often by averaging (for regression) or voting (for classification), to produce a final prediction [69].

3.5.3.1 RANDOM FOREST

Random Forest is an ensemble method, falling under bagging techniques, used for both classification and regression tasks. It constructs multiple decision trees (base learners) during training, using random sampling of observations and predictor variables from the training data.

A unique feature of Random Forest is the random selection of a subset of variables for splitting at each decision tree node, introducing diversity among the trees, and improving the ensemble’s performance. The final prediction is determined by a majority vote among the trees for classification tasks, or by averaging the predictions for regression tasks.

Random Forest is renowned for its ability to handle noisy data, prevent overfitting, and generalize well to unseen instances. Its robustness and accuracy, along with its ability to handle large datasets and require minimal hyperparameter tuning, contribute to its popularity. The randomness introduced in tree construction and variable selection enhances the ensemble's diversity, improving its resilience to outliers and overall performance. An important benefit of Random Forest is its insensitivity to the scale of predictor variables, enabling it to handle predictor variables with different scales without a significant impact on performance.

3.5.4 BOOSTING

Boosting also involves combining the outputs of multiple models, but the process is sequential and progresses over time. Base models are trained iteratively, with each subsequent model focusing on the observations that the previous models struggled with, assigning them higher weights. The final prediction is a weighted sum of the individual models, where the weights are determined by their performance.

3.5.4.1 ADABOOST (ADAPTIVE BOOSTING),

AdaBoost is an ensemble learning method primarily used for classification tasks. Unlike bagging which builds multiple models simultaneously, AdaBoost creates a sequence of weak learners iteratively. The primary objective of AdaBoost is to concentrate on the observations that the current model finds challenging.

The process begins with the training of a weak learner on the original dataset. A weak learner is essentially a classification algorithm that performs slightly better than random chance. The adaptability of AdaBoost is a critical concept, it assigns more weight to observations misclassified in the previous iteration, compelling the subsequent weak learners to focus on the challenging cases. In each iteration, a new weak learner is trained on the updated dataset, which now has increased weights on the misclassified observations. This iterative training process is repeated for a predefined number of iterations or until a perfect model is achieved.

Each weak learner is given a weight based on its accuracy in the training process. Higher accuracy results in a higher weight. The final prediction is made by combining the weighted votes of all weak learners, producing a more robust model than any individual weak learner.

AdaBoost is renowned for its ability to enhance the accuracy of weak models and its resistance to overfitting. It is adaptable to different types of weak learners. However, it can be sensitive to noisy data and outliers, and the performance may degrade if the weak learners are too complex or prone to overfitting.

3.5.4.2 GRADIENT BOOSTING MACHINE

Gradient Boosting Machine (GBM) is also an ensemble learning technique useful both for classification and regression tasks. It builds a robust model by combining multiple weak learners, much like AdaBoost. However, GBM differs in that it constructs trees sequentially, correcting errors made by previous trees by fitting a new tree to the residual errors of the combined ensemble. GBM starts with an initial model, typically a shallow decision tree, to the original data. This model serves as the first weak learner, and the residuals, or differences between predicted and actual values, are computed based on this model's predictions.

Subsequent trees are then constructed to predict the residuals from the previous step, minimizing the residuals left by the combined ensemble of the existing trees. A learning rate parameter controls the contribution of each tree to the final ensemble. A lower learning rate results in a more robust model but requires more trees to achieve similar performance, and to prevent overfitting, each tree is constructed with limited depth and pruned during training.

The final prediction is the sum of the forecasts from all the trees, reducing error and improving predictive accuracy. Gradient Boosting Machines, particularly implementations like Extreme Gradient Boosting (XGBoost), are popular due to their ability to handle complex relationships, provide high predictive accuracy, and handle variable interactions.

3.5.4.3 EXTREME GRADIENT BOOSTING (XGBOOST)

Extreme Gradient Boosting (XGBoost) is a highly efficient Gradient Boosting Machine implementation. It is popular in machine learning competitions and applications due to its exceptional performance.

XGBoost extends gradient boosting by incorporating regularization techniques, such as Lasso (L1) and Ridge (L2), to control overfitting. The model is trained by minimizing a user-specified loss function, combined with the regularization terms in the objective function. XGBoost optimizes the objective function by a gradient-based optimization algorithm.

Instead of growing an entire tree and pruning it later, XGBoost uses a depth-first approach for tree construction, considering the global structure of the tree. It supports parallel and distributed computing for efficacy and scalability, enabling faster training on large datasets across multiple processors.

XGBoost provides insights into variable importance, aiding variable selection and interpretation. It has a robust mechanism for handling missing values and retaining information from observations with missing data during training. It also supports early stopping to prevent overfitting when the model's performance on a validation dataset ceases to improve. XGBoost's combination of regularization, efficient tree construction, parallel processing, and advanced optimization techniques makes it a robust and widely adopted algorithm. Its versatility and ability to achieve state-of-the-art results across various domains have contributed to its popularity.

3.5.5 OTHER MACHINE-LEARNING METHODS

3.5.5.1 Support Vector Machines (SVM)

Support Vector Machines (SVM) is a machine learning algorithm used for classification and regression tasks [105]. It's effective in high-dimensional spaces and can handle multiclass classification problems. The primary goal of Support Vector Machines is to find a hyperplane that best separates data points into different classes or predicts the outcome variable.

SVM finds a hyperplane (the black dotted line in Figure 3.2 is the optimal hyperplane) that linearly separates the data into different classes (square is *Yes, disability pension,* and star is *No, disability pension*), maximizing the margin (the distance between the blue lines) between classes for a more robust model, as illustrated in Figure 3.2. It can handle non-linear relationships using the kernel trick, which transforms input variables into a higher-dimensional space. In scenarios where data is not perfectly separable, SVM introduces a soft margin concept, allowing some misclassification by introducing a penalty term for instances on the wrong side of the decision boundary. The regularization

parameter (C) controls the trade-off between a smooth decision boundary and correct classification. A minor C allows for a softer margin and more misclassifications, while a larger C enforces a stricter margin.
Support Vector Machines versatility, effectiveness in high-dimensional spaces, and ability to handle non-linear relationships through kernel functions make it successful in various domains. The choice of the kernel and tuning parameters significantly impacts SVM's performance.

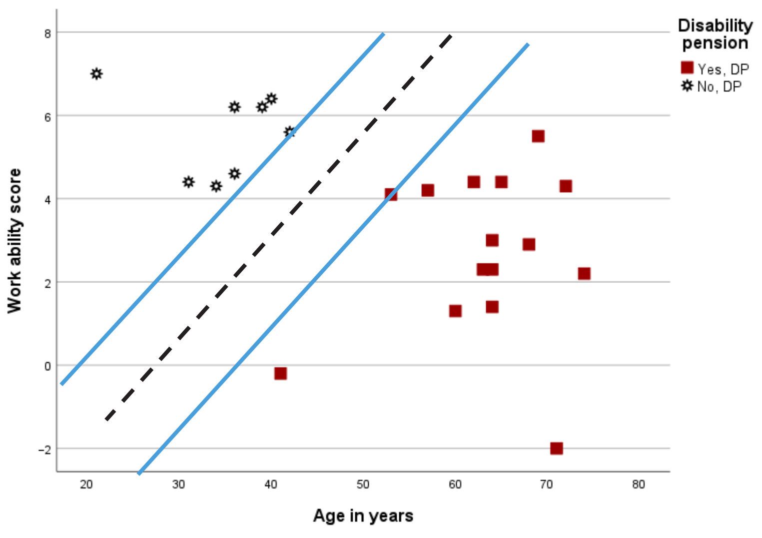

Figure 3.2: A simple illustration of Support Vector Machines using imaginary data, the optimal hyperplane (black dotted line) separating the data into disability pension (yes) and disability pension (no)

3.5.5.2 NAIVE BAYES CLASSIFIER

The Naive Bayes classifier is a probabilistic machine learning algorithm based on Bayes' theorem, used for classification tasks. It calculates the likelihood of a binary outcome (class) given predictor variables. The "Naive" assumption is that the variables used to describe an observation are conditionally independent given the class of an outcome.

The classifier estimates two types of probabilities: *Class Prior Probability*, the frequency of each class in the training data, and *Conditional Probability*, the likelihood of observing each predictor given the class.

Given a new instance, the classifier calculates the probability of each class given the observed variables using Bayes' theorem and the independence assumption. The class with the highest probability is assigned as the predicted class.

Variations of Naive Bayes classifiers exist for different data types, including Bernoulli, Gaussian, and Multinomial Naive Bayes. Despite its simplicity and "naive" assumptions, it is a robust and widely used algorithm, often performing well in practice. However, violations of the independence assumption can impact accuracy, and the model may be sensitive to irrelevant predictors.

3.6 EVALUATION OF PERFORMANCE

Prediction models for binary outcomes yield both a continuous probability of an outcome and the predicted outcome based on an arbitrary cut-off (such as 0.50). For instance, a predicted probability of disability pension ≥ 0.50 will categorize a worker as having disability pension. The worker with a predicted probability of disability pension at 0.51 is classified identically to one with a predicted probability of 0.99. Despite their different probabilities, both workers will be classified as receiving a disability pension. However, the probability estimate of 0.99 suggests a higher level of confidence in the classification as disability pension compared to the worker with a probability of 0.51.

Two fundamental aspects of model performance are discrimination and calibration, both essential for practical utility [106]. Discrimination measures the model's ability to distinguish between workers who are and are not awarded a disability pension. This is quantified by the c-statistic, which assesses the probability that a randomly selected individual with disability pension has a higher predicted probability than one without [66, 106]. A c-statistic of 1 indicates perfect discrimination, while 0.5 represents random predictions, akin to the area under a receiver operating characteristic (AUC) for binary outcomes.

On the other hand, calibration evaluates the accuracy of probabilistic predictions and ensures that predicted probabilities align well with actual probabilities. There is no single best method for measuring the calibration, certain methods are commonly used and have their own strengths and limitations. Calibration plots serve as a valuable tool for assessing the quality of class probabilities. These plots compare observed event rates with predicted class probabilities, with well-calibrated models demonstrating points along a 45-degree line. Such visualizations help evaluate the alignment between predicted probabilities and actual outcomes, thereby gauging the reliability of the classification model.

The Brier score, calculated as the mean squared prediction error, quantifies the average discrepancy between the predicted probabilities of the outcome and the actual observed outcomes. Rescaled to a range from 0% to 100%, with lower scores indicating better calibration, the Brier score provides insight into model accuracy [106]. Additionally, Spiegelhalter's z-test statistic assesses calibration by testing the hypothesis that the model is well-calibrated.

Other valuable metrics such as accuracy (proportion of correct true positive and true negative), precision (positive predictive value = (True Positives) / (True Positives + False Positives)), and recall (sensitivity = (True Positives) / (True Positives + False Negative), and F1 score (Harmonic Mean of Precision and Recall), rely on the distinction between true and false predicted outcomes made by the prognostic model, compared to actual outcomes. However, these can be influenced by the arbitrary cut-off used for class membership, based on specific objectives such as minimizing false negatives to detect illnesses in healthcare examples.

In Study IV, the predictive power of the techniques was evaluated using the c-statistic. The model's predictions were evaluated to determine its calibration, or the accuracy of its forecasts, by employing the scaled Brier score, an indicator that measures the average squared differences between predicted probabilities of disability pension and the observed outcome [106]. Additionally, Spiegelhalter's z-test statistic was computed to offer a further assessment of the model's calibration [106].

3.7 VALIDATION OF PREDICTION MODELS

In predictive modeling, internal and external validation are crucial for assessing the model's performance [66]. Validation generally implies that a prognostic model is not only effective for the data it was trained on, but also maintains its performance when applied to unseen data (validation set), which may differ from the training data used for its development.

Internal validation is often performed using *k-fold cross-validation or bootstrapping.* The *k-fold cross-validation* involves randomly dividing the data into *k* subsets of equal size [103]. The model is then trained on *k-1* subsets and evaluated on the remaining subset. This process is repeated *k* times, with each subset serving as the validation dataset once. The results are then averaged to estimate the model's performance. A further step is to repeat the entire *k-fold cross-validation* process say 5 times. In 5 repeats of *k-fold cross-validation*, the entire *k-fold cross-validation* process is completed 5 times. The performance measures obtained from each round of *k-fold cross-validation*, which amounts to a total of 5**k* (5 repeats * *k* folds), will be aggregated to yield a reliable performance estimate. This technique enhances the stability and robustness of the model's performance estimate, reducing the variability associated with a single trial of *k-fold cross-validation*.

On the other hand, **external validation** is crucial to ensure the generalizability of a prediction model. There are several methods for external validation, such as:

Temporal validation: The model is validated on data from different times. For example, we might validate a model on the most recent data, this would qualify as a temporal external validation.

Geographical validation: The model is validated on data from different geographical locations than the one it was trained on. This helps to assess the model's performance across diverse populations and environments.

Independent researcher validation: The model is validated by researchers who were not involved in its development. This can help to eliminate potential biases and ensure objectivity.

Together, internal, and external validation provide a comprehensive evaluation of a prediction model's performance and generalizability.

3.8 ANALYTICAL METHODS USED

In Study I, log-binomial regressions were employed to compute risk differences and relative risks (RRs) along with their respective 95% confidence intervals. The relative risk, RR, with a 95% CI, estimated the association between having a specific syndrome (yes/no) and binary disability pension. Besides the unadjusted models, multivariable log-binomial regressions adjusting for age (in years) were also performed.

In Study II, a counterfactual-based mediation analysis approach was utilized to calculate the effects (total, direct, and indirect) on the risk ratio scale (adjusted relative risks) (RRs) and proportion of effect mediated (PM, on the risk difference scale) along with their respective bias-corrected bootstrap 95% confidence intervals (CI) [92]. Drawing upon existing research and subject-matter expertise, a causal model was meticulously constructed by integrating the Cancer and Work model (Figure 1.2). This causal model, depicted in Figure 3.3, included several confounding variables, including age at baseline (in years), marital status at baseline, socioeconomic group, and number of comorbidities.

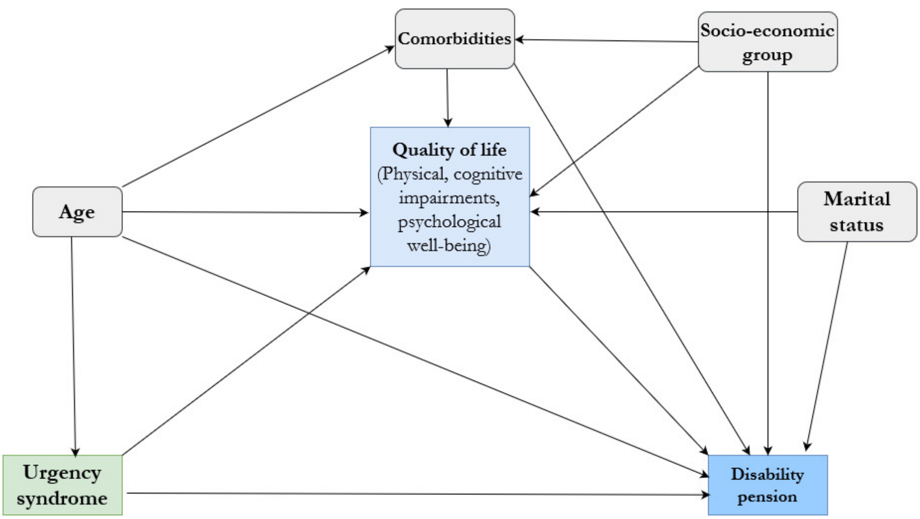

Figure 3.3: The causal mediation model used in Study II with confounders measured at baseline.

In Study III, ordinary linear regression-based mediation analysis was applied. Considering existing research findings, a theoretical model incorporating the following four mediating variables of interest was initially considered: demands at work affecting private life (work–home interference), feelings of control over private life (control over private life), feeling well-rested upon waking (recovery), and physical activity in leisure time (physical activity). Theoretical model was transformed into a mediation model by incorporating mediators into the statistical framework. But physical activity was removed as it showed non-significant mediation relationships during the sequential building of the final mediation model. The PROCESS macro in IBM SPSS Statistics for Windows was utilized to calculate the regression coefficients of total, direct, and indirect effects, the proportion of effect-mediated, and their 95% confidence intervals (CI) [107].

The level of significance (alpha) was set at 5%. All hypothesis tests were two-sided. All statistical analyses were performed using SAS 9.4 statistical analysis software (SAS Institute, RRID: SCR_008567), IBM SPSS Statistics for Windows (IBM Corp, RRID: SCR_002865), and RStudio v2022.07.2 Build 576 “Spotted Wakerobin”, RRID: SCR_000432).

3.8.1 HYPERPARAMETER TUNING AND CALIBRATION IN STUDY IV

In Study IV, we utilized logistic regression, AdaBoost, Extreme Gradient Boosting, Gradient Boosting Machine, Naïve Bayes classifier, Random Forest, and Support Vector Machine to estimate the risk of being awarded a disability pension. The associations between disability pension and predictors were initially computed using the logistic regression for each model. Afterward, a set of candidate values of hyperparameters was defined for tuning. For each candidate set, five repeats of *10-fold cross-validation* were employed to adjust the hyperparameters. The objective was to maximize the area under the receiver operating curve (AUC), which served as an estimate for these hyperparameters [103].

The combination of hyperparameters resulting in the most significant AUC for each machine learning algorithm was determined. These optimal hyperparameters were then utilized to develop predictive models using training data from 2011. The validation data from 2013 was predicted using these predictive models, and the probabilities of disability pension were estimated while recording the model's performance metrics, c statistic (AUC), and scaled Brier score.

To ensure the probabilities closely reflect the likelihood of disability pension in the test set, these probabilities were further calibrated by constructing additional logistic regression models [103]. The initial probabilities were used as input, with disability pension as the output. The caret package in R was employed for the tuning process across all techniques [108].

4 RESULTS

The findings derived from the four studies are presented in this section. Readers are advised to refer to the respective papers for a more comprehensive presentation of the results.

4.1 RESULTS STUDY I-II

4.1.1 COHORT CHARACTERISTICS

From the initial pool, a cohort of 247 gynecological cancer survivors met the eligibility criteria for Studies I and II. The primary exclusion criterion being age, with only individuals aged between 19 and 64 years being eligible for disability pension benefits [47]. Some occupational pensions are automatically disbursed at the age of 65 [47].

In the cohort of gynecological cancer survivors, the most common diagnosis was endometrial cancer, accounting for 42% of diagnoses, followed by cervical (38%), ovarian (9%), fallopian tube (2%), vaginal (4%), sarcoma uteri (4%), and vulvar (1%) cancers. Radiation-induced gastrointestinal syndromes were prevalent among the cohort. Specifically, 47% (116 out of the 247) survivors exhibited at least one syndrome, while the remaining 53% (131 out of 247) did not manifest any syndromes. In the cohort, 37% (91 females) presented with urgency syndrome at baseline. The distribution of survivors across specific syndromes varied, as outlined in Table 4.1, representing the baseline clinical and background data for all gynecological cancer survivors and gynecological cancer survivors with and without syndromes.

Table 4.1: Baseline (in 2006) clinical and demographic data for all gynecological cancer survivors and gynecological cancer survivors with and without syndromes (Counts (percentages))

	All survivors	Urgency syndrome		Leakage syndrome		Blood Discharge syndrome	
		Yes	No	Yes	No	Yes	No
	n=247	n=91 (37 %) [a]	n=156 (63 %) [a]	n=77 (31%) [a]	n=170 (69 %) [a]	n=31 (13 %) [a]	n=216 (87 %) [a]
Age							
16-29 years	2 (1 %)	2 (2 %)	0	0	2 (1 %)	0	2 (1 %)
30-49 years	66 (27 %)	24 (26 %)	42 (27 %)	19 (25 %)	47 (28 %)	7 (23 %)	59 (27 %)
50-64 years	179 (72 %)	65 (71 %)	114 (73 %)	58 (75 %)	121 (71 %)	24 (77 %)	155 (72 %)
Marital status							
Married/living with partner	160 (65 %)	58 (64 %)	102 (66 %)	46 (60 %)	114 (67 %)	16 (52 %)	144 (67 %)
Widow	12 (5 %)	5 (5 %)	7 (5 %)	6 (8 %)	6 (4 %)	3 (10 %)	9 (9 %)
Has partner but lives alone	21 (9 %)	11 (12 %)	10 (6 %)	12 (16 %)	9 (5 %)	2 (6 %)	19 (9 %)
Single	53 (22 %)	17 (19 %)	36 (23 %)	13 (17 %)	40 (24 %)	12 (32 %)	43 (20 %)
Not stated	1		1		1		1
Self-reported employment							
Employed	169 (69 %)	54 (59 %)	115 (75 %)	43 (57 %)	126 (75 %)	16 (53 %)	153 (72 %)
Disability pension	35 (14 %)	22 (24 %)	13 (9 %)	20 (26 %)	15 (9 %)	9 (30 %)	26 (12 %)
Unemployed	12 (5%)	3 (3 %)	9 (6 %)	4 (5 %)	8 (5 %)	2 (7 %)	10 (5 %)
Housewife, other	10 (4 %)	3 (3 %)	7 (5 %)	5 (7 %)	5 (3 %)	1 (3 %)	9 (4 %)
Sickness absence	9 (4 %)	7 (8 %)	2 (1 %)	1 (1 %)	8 (5 %)	1 (3 %)	8 (4 %)
Student	5 (2 %)	2 (2 %)	3 (2 %)	2 (3 %)	3 (2 %)	1 (3 %)	4 (2 %)
Retired	4 (2 %)	0	4 (3 %)	1 (1 %)	3 (2 %)	0	4 (2 %)
Not stated	3		3	1	2	1	2
Diagnosis							
Endometrial cancer	104 (42 %)	38 (42 %)	66 (42 %)	27 (35 %)	77 (45 %)	9 (29 %)	95 (44 %)
Cervical cancer	93 (38 %)	33 (36 %)	60 (38 %)	27 (35 %)	66 (39 %)	11 (35 %)	82 (38 %)
Ovarian cancer	23 (9 %)	10 (11 %)	13 (8 %)	11 (14 %)	12 (7 %)	4 (13 %)	19 (9 %)
Vaginal cancer	11 (4 %)	4 (4 %)	7 (4 %)	5 (6 %)	6 (4 %)	5 (16 %)	6 (3 %)
Sarcoma uteri	9 (4 %)	3 (3 %)	6 (4 %)	3 (4 %)	6 (4 %)	0	9 (4 %)
Fallopian tube cancer	5 (2 %)	2 (2 %)	3 (2 %)	3 (4 %)	2 (1 %)	1 (3 %)	4 (2 %)
Vulvar cancer	2 (1 %)	1 (1 %)	1 (1 %)	1 (1 %)	1 (1 %)	1 (3 %)	1 (1 %)
Treatment modality							
Surgery + EBRT[c] + BT[d]	110 (45 %)	35 (38 %)	75 (48 %)	24 (31 %)	86 (51 %)	9 (29 %)	101 (47 %)
Surgery + EBRT[c] + BT[d] + Chemo[e]	56 (23 %)	18 (20 %)	38 (25 %)	15 (19 %)	41 (24 %)	6 (19 %)	50 (23 %)
Surgery + EBRT[c] + Chemo[e]	28 (11 %)	14 (15 %)	14 (9 %)	15 (19 %)	13 (8 %)	5 (16 %)	23 (11 %)
Surgery + EBRT[c]	18 (7 %)	8 (9 %)	10 (6 %)	7 (9 %)	11 (7 %)	5 (16 %)	13 (6 %)
EBRT[c] + BT[d] + Chemo[e]	17 (7 %)	5 (5 %)	12 (8 %)	4 (5 %)	13 (8 %)	2 (6 %)	15 (7 %)
EBRT[c] + BT[d]	11 (4 %)	5 (5 %)	6 (4 %)	6 (8 %)	5 (3 %)	1 (3 %)	10 (5 %)
EBRT[c] + Chemo[e]	5 (2 %)	5 (5 %)	0	5 (6 %)	0	2 (6 %)	3 (1 %)
EBRT[c]	1 (<1 %)	1 (1 %)	0	1 (1 %)	0	1 (3 %)	0
Not stated	1		1				1
Parity							
Never given birth	80 (32 %)	27 (30 %)	53 (34 %)	21 (27 %)	59 (35 %)	8 (26 %)	72 (33 %)
1–3 Children	151 (61 %)	56 (62 %)	95 (61 %)	52 (68 %)	99 (58 %)	19 (61 %)	132 (61 %)
> 3 Children	16 (6 %)	8 (9 %)	8 (5 %)	4 (5 %)	12 (7 %)	4 (13 %)	12 (6 %)

Abbreviations: BT, brachy therapy; chemo, chemotherapy; EBRT, external beam radiation therapy.
[a]Counts(percentage) within each category of syndrome
[b]Counts and percentage of survivors in each category

Continued Table 4.1: Baseline (in 2006) clinical and demographic data for gynecological cancer survivors with and without syndromes (Counts (percentages))

	Gynecological survivors with			
	No syndrome	One syndrome	Two syndromes	Three syndromes
	n = 131 (53 %) [a]	n = 50 (20%) [a]	n = 49 (20%) [a]	n = 17 (7 %) [a]
Age				
16-29 years	0	2 (4 %)	0	0
30-49 years	36 (27 %)	13 (26 %)	14 (29 %)	3 (18 %)
50-64 years	95 (73 %)	35 (70 %)	35 (71 %)	14 (82 %)
Marital status				
Married/living with partner	86 (66 %)	38 (76 %)	26 (53 %)	10 (59 %)
Widow	6 (5 %)	0 (0 %)	4 (8 %)	2 (12 %)
Has partner but lives alone	6 (5 %)	6 (12 %)	8 (16 %)	1 (6 %)
Single	32 (25 %)	6 (12 %)	11 (22 %)	4 (24 %)
Not stated	1			
Self-reported employment				
Employed	100 (77 %)	33 (69 %)	28 (57 %)	8 (47 %)
Disability pension	10 (8 %)	5 (10 %)	14 (29 %)	6 (35 %)
Unemployed	7 (5 %)	2 (4 %)	2 (4 %)	1 (6 %)
Housewife, other	5 (4 %)	2 (4 %)	2 (4 %)	1 (6 %)
Sickness absence	2 (2 %)	5 (10 %)	2 (4 %)	0
Student	3 (2 %)	0	1 (2 %)	1 (6 %)
Retired	3 (2 %)	1 (2 %)	0	0
Not stated	1	2	0	0
Diagnosis				
Endometrial cancer	57 (44 %)	24 (48 %)	19 (39 %)	4 (24 %)
Cervical cancer	52 (40 %)	17 (34 %)	18 (37 %)	6 (35 %)
Ovarian cancer	10 (8 %)	3 (6 %)	8 (16 %)	2 (12 %)
Vaginal cancer	4 (3 %)	3 (6 %)	1 (2 %)	3 (18 %)
Sarcoma uteri	5 (4 %)	2 (4 %)	2 (4 %)	0
Fallopian tube cancer	2 (2 %)	1 (2 %)	1 (2 %)	1 (6 %)
Vulvar cancer	1 (1 %)	0	0	1 (6 %)
Treatment modality				
Surgery + EBRT[c] + BT[d]	66 (51 %)	23 (46 %)	18 (37 %)	3 (18 %)
Surgery + EBRT[c] + BT[d] + Chemo[e]	30 (23 %)	17 (34 %)	5 (10 %)	4 (24 %)
Surgery + EBRT[c] + Chemo[e]	11 (8 %)	4 (8 %)	9 (18 %)	4 (24 %)
Surgery + EBRT[c]	7 (5 %)	4 (8 %)	5 (10 %)	2 (12 %)
EBRT[c] + BT[d] + Chemo[e]	11 (8 %)	1 (2 %)	5 (10 %)	0
EBRT[c] + BT[d]	5 (4 %)	1 (2 %)	4 (8 %)	1 (6 %)
EBRT[c] + Chemo[e]	0	0	3 (6 %)	2 (12 %)
EBRT[c]	0	0	0	1 (6 %)
Not stated	1			
Parity				
Never given birth	47 (36 %)	14 (28 %)	15 (31 %)	4 (24 %)
1–3 Children	77 (59 %)	32 (64 %)	31 (63 %)	11 (65 %)
> 3 Children	7 (5 %)	4 (8 %)	3 (6 %)	2 (12 %)

Abbreviations: BT, brachy therapy; chemo, chemotherapy; EBRT, external beam radiation therapy.

[a]Counts(percentage) within each category of syndrome

[b]Counts and percentage of survivors in each category

4.1.2 STUDY I

At the 2-year follow-up, disability pension benefits had been granted to 27% of gynecological cancer survivors. The proportion of recipients varied across the five syndromes and the cumulative syndromes (number of syndromes), as illustrated in Figure 4.1.

Notably, among survivors with urgency syndrome, 38% were recipients of a disability pension, in contrast to only 20% of survivors without urgency syndrome. This disparity resulted in a relative risk (RR) of 1.9 (95% CI 1.3–2.9). Similarly, survivors with blood discharge syndrome and leakage syndrome had an RR of 2.0 (95% CI 1.3–3.2) and 2.1 (95% CI 1.4–3.1), respectively. Interestingly, survivors with excessive gas discharge syndrome exhibited a noteworthy RR, although the 95% CIs included the null value. The RR for disability pension among survivors with excessive mucus discharge syndrome was 1.0, with the 95% CI encompassing the null value but extending to potentially pertinent values (see Figure 4.1).

Syndromes	Disability pension No. (%)	Relative risk (95% CI)
Sum of syndromes		
Three syndromes	9 (53 %)	2.9 (1.6, 5.1)
Two syndromes	22 (45 %)	2.5 (1.5, 3.9)
One syndrome	11 (22 %)	1.2 (0.6, 2.3)
None	24 (18 %)	reference
Urgency syndrome		
Yes	35 (38 %)	1.9 (1.3, 2.9)
No	31 (20 %)	reference
Leakage syndrome		
Yes	32 (42 %)	2.1 (1.4, 3.1)
No	34 (20 %)	reference
Blood discharge		
Yes	15 (48 %)	2.0 (1.3, 3.2)
No	51 (24 %)	reference
Excessive gas discharge		
Yes	20 (36 %)	1.5 (1.0, 2.3)
No	46 (24 %)	reference
Excessive mucus discharge		
Yes	15 (27 %)	1.0 (0.6, 1.6)
No	51 (27 %)	reference

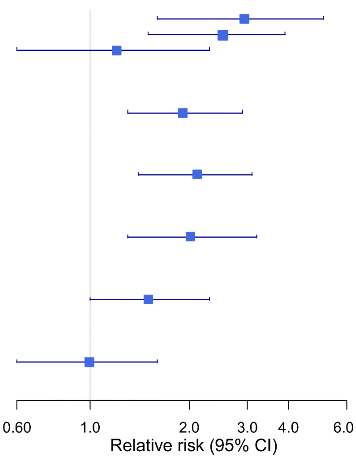

Figure 4.1: Counts (percentage) and relative risk (RR) (95% confidence interval) of receiving disability pension at 2-year follow-up, RR of > 1 indicates harm. RR was obtained from log-binomial regression.

Moreover, survivors with one or more syndromes exhibited a higher RR of disability pension compared to those without any syndromes. At the 2-year follow-up, survivors with two syndromes had an RR of 2.5 (95% CI 1.5–3.9), while those with three syndromes had an RR of 2.9 (95% CI 1.6–5.1) compared to syndrome-free survivors (Figure 4.1). Remarkably, a monotonic increase in effect measures was observed. Additionally, the Cochran-Armitage test, which treated the sum of syndromes as an ordinal predictor and disability pension as a binary outcome, was statistically significant ($p<0.05$), supporting the trend hypothesis.

Multivariable regression analysis indicated that age (in years) did not alter the interpretation regarding the association between a syndrome and the likelihood of receiving a disability pension, as presented in Appendix 1.

Gynecological cancer survivors with specific syndromes were between 1.9 and 2.9 times more likely to receive a disability pension compared to those without such syndromes, even after adjusting for age in multivariable regression. The risk of disability pension was particularly heightened among survivors with three syndromes and was next highest among those with blood discharge syndrome (Figure 4.1).

4.1.3 STUDY II

Baseline clinical and background data are outlined in Table 4.1, while Table 4.2 provides descriptive statistics for mediator variables among all 247 gynecological survivors.

Interactions between urgency syndrome and mediators were noticeable across various quality of life (QoL) dimensions. These include global QoL, physical health, physical strength, psychological well-being, and satisfaction with sleep. The incorporation of exposure–mediator interactions amplified the indirect effects for these mediators, as shown in Table 4.3. Almost all adjusted total effects (TEs), nearly half of the adjusted direct effects (NDEs), and the indirect effects (NIEs), and their 95% confidence intervals (CIs) were greater than 1. This indicates an increased likelihood of receiving a disability pension. Comparable magnitudes were observed across all investigated mediators, see Table 4.3, with similar patterns noted for self-reported physical and psychological QoL aspects. Some QoL dimensions exhibited higher values, as shown in Table 4.3.

Table 4.2: Descriptive statistics for the mediator variables, physical and psychological aspects of quality of life (QoL), for all gynecological cancer survivors

Self-assessed quality of life		Median (1^{st}–3^{rd} quartile)	Counts (%) [a]
Global quality of life	Low to moderate	5 (4–6)	140 (57)
Physical aspects			
Global physical health	Low to moderate	5 (4–6)	167 (68)
Physical strength (condition)	Low to moderate	5 (4–6)	165 (67)
Psychological aspects			
Satisfied with sleep	Low to moderate	4 (3–6)	123 (50)
Satisfied with concentration	Low to moderate	5 (4–6)	152 (62)
Satisfied with memory	Low to moderate	5 (4–6)	149 (61)
Psychological wellbeing	Low to moderate	5 (4–6)	167 (68)
Self-esteem	Low to moderate	5 (4–6)	146 (59)
Having meaning in life	Low to moderate	6 (4–7)	175 (71)
Worry or anxiety	High [b]	3 (2–5)	105 (43) [b]
Depressed or feeling sad	High [b]	3 (2–5)	105 (43) [b]

[a] This shows the frequency and the percentage of survivors scoring ≤ median.
[b] For worry or anxiety and depressed or feeling sad, the frequency and the percentage of those reporting values are higher than the median.

For instance, in a single mediation analysis, approximately 46% of the TE (RR = 2.3 (1.1–3.7)) of urgency syndrome on a disability pension was mediated through global physical health, although the 95% CI for the proportion mediated ranged from 22% to 110%. A more significant proportion of the effect was mediated via global and physical aspects of QoL (35%–71%) compared with psychological elements (2%–47%) Table 4.3.

Sensitivity analysis using mediational E-values indicated significant roles of most mediators in explaining the association, mitigating the potential impact of unmeasured confounding (Table 4.4). The likelihood of an unmeasured confounder explaining the association, instead of the mediator, was deemed low except for satisfaction with memory, meaning in life, worry or anxiety, and feeling depressed/sad. They all exhibited an E-value less than 1.6, see Table 4.4. In an additional sensitivity analysis, after excluding gynecological survivors who received disability pension during 2004–2006, the sample size reduced to n = 187, with only seven cases of disability pension during follow-up in 2008 (Table S3 in the supplementary material of the published Study II). Insufficient events (seven instances) limited the analysis to detect effect sizes of interest (RR $\geq$ 2), although indications from both mediation analyses (lenient inclusion criterion with n = 247 and rigorous inclusion criterion with n = 187) pointed in the same direction. Despite almost identical indirect effects, the latter mediation analysis exhibited more robust total and direct effects, albeit with abnormally wide CIs returned by the statistical toolbox PROC CAUSALMED. For further details, refer to Table S3 in the supplementary material of the published Study II.

Table 4.3: Adjusted NDE, NIE, and TE of radiation-induced urgency syndrome on disability pension in the presence of exposure–mediator interaction (n = 247 gynecological cancer survivors). Data on disability pension were obtained from the official register; the table shows the findings of mediation analysis while adjusting for age (in years), marital status, occupation-based socioeconomic group, and number of comorbidities.

Mediator (M)	Adjusted relative risk[a] (95% CI)			Proportion mediated[b] (95% CI)
	Natural direct effect (NDE)	Natural indirect effect (NIE)	Total effect (TE)	
Global quality of life	1.3 (0.8–2.4)	**1.6 (1.3–2.0)**	**2.1 (1.2–3.7)**	**71% (40%–163%)**
Physical aspects				
Global physical health	1.7 (0.8–2.8)	**1.3 (1.0–1.6)**	**2.3 (1.1–3.7)**	**46% (22%–110%)**
Physical strength (condition)	1.7 (0.9–3.4)	**1.2 (1.0–1.4)**[c]	**2.1 (1.2–3.9)**	**35% (7%–64%)** [c]
Psychological aspects				
Satisfied with sleep	**1.6 (0.8–3.1)**	**1.3 (1.1–1.7)**	**2.2 (1.1–4.1)**	**47% (18%–111%)**
Psychological wellbeing	**1.7 (1.0–2.5)** [c]	**1.3 (1.1–1.6)**	**2.3 (1.4–3.2)** [c]	**43% (15%–71%)** [c]
Satisfied with concentration	**1.5 (1.0–2.3)** [c]	1.2 (1.0–1.4) [c]	**1.8 (1.0–2.6)**	**36% (2%–70%)**
Self-esteem	**1.9 (1.1–5.8)**	1.2 (1.0–1.3) [c]	**2.2 (1.3–5.0)**	**27% (1%–52%)**
Worry or anxiety	1.6 (0.8–2.7)	1.1 (0.9–1.4)	1.8 (0.9–2.9)	20 (-13%–103%)
Having meaning in life	1.7 (0.9–3.7)	1.1 (1.0–1.3)	**1.8 (1.0–4.0)**	20% (-12%–93%)
Satisfied with memory	**1.9 (1.0–2.8)**	1.1 (0.8–1.3)	**2.0 (1.1–2.7)**	12% (-32%–49%)
Depressed or feeling sad	**2.2 (1.3–4.5)**	1.0 (0.9–1.3)	**2.3 (1.4–4.5)**	2% (-17%–74%)

Abbreviations: CI, confidence interval; *Note:* Estimates of natural direct effect, indirect effect, total effect, and proportion of effect mediated were obtained by using an aspect of self-assessed QoL as a mediator(M) for the association between urgency syndrome(X) and disability pension(Y) in presence of XM-interaction. **Bold numbers** indicate a statistically significant effect at a 5% level of significance. NDE = the contrast between the counterfactual outcome while being exposed and the counterfactual outcome while the same individual was not exposed, mediator assuming the value it would have taken while being not exposed. NIE = the contrast, having set the exposure = Yes, exposed between the counterfactual outcome (mediator assumed whatever value it would have taken at a value of the exposure = Yes, exposed and the counterfactual outcome if the mediator assumed whatever value, it would have taken at a reference value of the exposure = not exposed).

[a] Adjusted relative risk with bootstrap bias-corrected 95% CI. [b] Proportion mediated = NDE*(NIE-1)/(NDE*NIE-1), with bootstrap bias corrected 95% CI. [c] Wald 95% CI.

Table 4.4: Sensitivity analysis of an unmeasured confounding using mediational E-values.

Mediator (M)	Adjusted relative risk[a] (95% CI) Natural direct effect (NDE)	Natural indirect effect (NIE)	Total effect (TE)	Proportion mediated[b] (95% CI)
Global quality of life	1.3 (0.8–2.4)	**1.6 (1.3–2.0)**	**2.1 (1.2–3.7)**	**71% (40%–163%)**
Physical aspects				
Global physical health	1.7 (0.8–2.8)	**1.3 (1.0–1.6)**	**2.3 (1.1–3.7)**	**46% (22%–110%)**
Physical strength (condition)	1.7 (0.9–3.4)	**1.2 (1.0–1.4)**[c]	**2.1 (1.2–3.9)**	**35% (7%–64%)** [c]
Psychological aspects				
Satisfied with sleep	**1.6 (0.8–3.1)**	**1.3 (1.1–1.7)**	**2.2 (1.1–4.1)**	**47% (18%–111%)**
Psychological wellbeing	**1.7 (1.0–2.5)** [c]	**1.3 (1.1–1.6)**	**2.3 (1.4–3.2)** [c]	**43% (15%–71%)** [c]
Satisfied with concentration	**1.5 (1.0–2.3)** [c]	1.2 (1.0–1.4) [c]	**1.8 (1.0–2.6)**	**36% (2%–70%)**
Self-esteem	**1.9 (1.1–5.8)**	1.2 (1.0–1.3) [c]	**2.2 (1.3–5.0)**	**27% (1%–52%)**
Worry or anxiety	1.6 (0.8–2.7)	1.1 (0.9–1.4)	1.8 (0.9–2.9)	20 (-13%–103%)
Having meaning in life	1.7 (0.9–3.7)	1.1 (1.0–1.3)	**1.8 (1.0–4.0)**	20% (-12%–93%)
Satisfied with memory	**1.9 (1.0–2.8)**	1.1 (0.8–1.3)	**2.0 (1.1–2.7)**	12% (-32%–49%)
Depressed or feeling sad	**2.2 (1.3–4.5)**	1.0 (0.9–1.3)	**2.3 (1.4–4.5)**	2% (-17%–74%)

Abbreviations: CI, confidence interval; *Note:* Estimates of natural direct effect, indirect effect, total effect, and proportion of effect mediated were obtained by using an aspect of self-assessed QoL as a mediator(M) for the association between urgency syndrome(X) and disability pension(Y) in presence of XM-interaction. **Bold numbers** indicate a statistically significant effect at a 5% level of significance. NDE = the contrast between the counterfactual outcome while being exposed and the counterfactual outcome while the same individual was not exposed, mediator assuming the value it would have taken while being not exposed. NIE = the contrast, having set the exposure = Yes, exposed between the counterfactual outcome (mediator assumed whatever value it would have taken at a value of the exposure = Yes, exposed and the counterfactual outcome if the mediator assumed whatever value, it would have taken at a reference value of the exposure = not exposed).

[a] Adjusted relative risk with bootstrap bias-corrected 95% CI. [b] Proportion mediated = NDE*(NIE-1)/(NDE*NIE-1), with bootstrap bias corrected 95% CI. [c] Wald 95% CI.

4.2 STUDY III

The study comprised 1432 participants who completed the questionnaires in 2012 and 2017 and reported stress and Work Ability Score measures at both time points. Refer to Appendix 2 for the flowchart depicting participant selection.

Among the sample, 60% were female (n = 857), while 40% were males (n = 575). The mean WAS at baseline (2012) was 8.2 (standard deviation = 1.7), which decreased to 7.9 (standard deviation = 1.8) after five years, Table 4.5.

All single mediation relationships were statistically significant, except for the indirect effects of physical activity in 2017. The percentages of the indirect impact of single mediation models ranged from 2% to 32%. Stress in 2017 emerged as a robust mediator between stress in 2012 and WAS in 2017, mediating 66% of this relationship.

Building upon the single mediation model, four mediators (all measured in 2017) were sequentially added, leading to a final multiple mediation model with *stress*, *work demands affecting private life*, *feelings of control over private life*, and *feeling well-rested upon waking up* mediating the stress in 2012-WAS association, see Figure 4.2 and Table 4.6. Additional details are provided in the published article.

In a concurrent analysis, performing a parallel quadruple mediation analysis, the mediating effect of each mediator was examined. All indirect effects were significant. The effect via pathway a_1b_1 accounted for 41% of the total effect, while a_2b_2, a_3b_3, and a_4b_4 contributed 14%, 19%, and 6% respectively, highlighting feelings of control over private life as a notable mediator, see Figure 4.3.

Combined, the four mediators explained 79% of the longitudinal relationship between stress in 2012 and work ability in 2017 for females and 87% for males. Notably, the direct effect (c′) became non-significant after incorporating these mediators in both genders. The detailed results of the gender-stratified quadruple mediation analysis in parallel are presented in Table 4.7. Small gender differences in mediating pathways are evident, with males exhibiting a slightly larger combined mediating effect. Additional details are provided in the published article.

Table 4.5: Descriptive statistics of the study population that answered the questions on stress and work ability in 2012 and 2017 (N = 1432). Abbreviation, std. dev = standard deviation.

	2012 Mean (std. dev)	**2017 Mean (std. dev)**
Age	27.1 (1.4)	32.1 (1.4)
Weight (in kg)	73.0 (14.5)	
Height (in cm)	173.3 (9.4)	
Work ability score (WAS)	8.2 (1.7)	7.9 (1.8)

	% (Counts)	**% (Counts)**
Highest completed education		
Primary school (9 years)	1 (16)	1 (14)
High School (12 years)	37 (531)	25 (356)
University or tertiary, <3 years	12 (169)	12 (168)
University or tertiary, >3 years	50 (715)	62 (890)
Work, Study		
Work or internship	73 (1050)	87 (1247)
Study	12 (172)	4 (56)
Both	15 (210)	9 (129)
Type of employment		
Permanent employment	68 (854)	88 (1208)
Trial employment	5 (59)	3 (35)
Temporary employment (hourly)	9 (115)	3 (37)
Limited-time (seasonal or project-based)	18 (224)	7 (94)
Family situation		
Married, partnership, cohabiting	61 (869)	75 (1073)
Girlfriend, boyfriend	12 (172)	6 (90)
Single	27 (387)	19 (265)
Stress		
Not at all	15 (217)	13 (186)
Just a little	33 (473)	31 (440)
To some extent	28 (405)	30 (426)
Pretty much	18 (262)	20 (284)

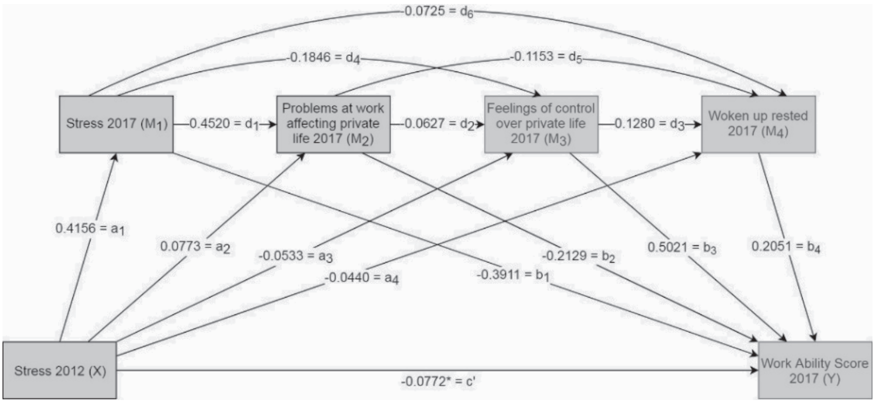

Figure 4.2: Direct acyclic graph representing the mediation model, also illustrating the effect sizes for each path. The effect sizes were obtained from **quadruple mediation model in series** utilizing ordinary least squares regressions.

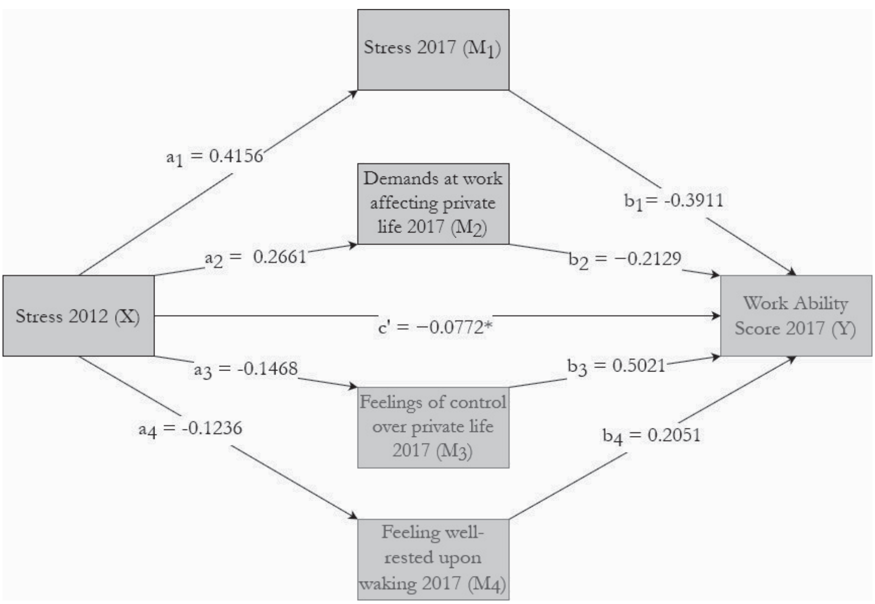

Figure 4.3: Direct acyclic graph representing the mediation model, also illustrating the effect sizes for each path. The effect sizes were obtained from **quadruple mediation in parallel** (without d-paths) utilizing ordinary least squares regressions.

Table 4.6: Total, direct, and indirect effects with 95% bias-corrected confidence interval. Results were obtained from **quadruple mediation analysis in series** utilizing ordinary least squares regressions. (Figure 4.2).

	Pathway	Effect	95% CI	Percent mediated via the pathway
Total effect	c	−0.3955	[−0.4764, −0.3146]	100
Direct effect	c′	−0.0772	[−0.1568, 0.0023]^{NS}	19.5
Total indirect effect	the sum of all indirect paths	−0.3182	[−0.3750, −0.2642]	80.5
Indirect effect 1	a_1b_1	−0.1626	[−0.2075, −0.1205]	41.1
Indirect effect 2	a_2b_2	−0.0165	[−0.0308, −0.0048]	4.2
Indirect effect 3	a_3b_3	−0.0268	[−0.0469, −0.0089]	6.8
Indirect effect 4	a_4b_4	−0.009	[−0.0195, −0.0006]	2.3
Indirect effect 5	$a_1d_1b_2$	−0.0402	[−0.0592, −0.0235]	10.2
Indirect effect 6	$a_1d_4b_3$	−0.0385	[−0.0544, −0.0251]	9.7
Indirect effect 7	$a_1d_6b_4$	−0.0062	[−0.0123, −0.0017]	1.6
Indirect effect 8	$a_2d_2b_3$	−0.0024	[−0.0052, −0.0005]	0.6
Indirect effect 9	$a_2d_5b_4$	−0.0018	[−0.0039, −0.0004]	0.5
Indirect effect 10	$a_3d_3b_4$	−0.0014	[−0.0031, −0.0003]	0.4
Indirect effect 11	$a_1d_1d_2b_3$	−0.0059	[−0.0104, −0.0021]	1.5
Indirect effect 12	$a_1d_1d_5b_4$	−0.0045	[−0.0076, −0.0019]	1.1
Indirect effect 13	$a_1d_4d_3b_4$	−0.002	[−0.0039, −0.0007]	0.5
Indirect effect 14	$a_2d_2d_3b_4$	−0.0001	[−0.0003, 0.0000]^{NS}	0.0
Indirect effect 15	$a_1d_1d_2d_3b_4$	−0.0003	[−0.0007, −0.0001]	0.1

Abbreviations: CI, confidence interval. NS, not statistically significant effect at the 5% level of significance
Note: Estimates of direct effect, indirect effect, total effect, and proportion of effect mediated were obtained from quadruple mediation analysis in series.
Total indirect effect was the sum of all indirect paths a1b1, a2b2, a3b3, a4b4, a1d1b2, a1d4b3, a1d6b4, a2d2b3, a2d5b4, a3d3b4, a1d1d2b3, a1d1d5b4, a1d4d3b4, a2d2d3b4, and a1d1d2d3b4.

Table 4.7: Gender-specific **quadruple mediation analysis in parallel**: total, direct, and indirect effects with 95% confidence intervals (CI)

Pathway	Effect	95% CI	Percent mediated via the pathway
Females (n = 857)			
Total effect (c)	−0.3577	[−0.4677, −0.2477]	100
Direct effect (c′)	−0.0760	[−0.1861, 0.0341]^{NS}	21.2
Total indirect effect	−0.2817	[−0.3547, −0,2154]	78.8
a_1b_1	−0.1404	[−0.1991, −0.0839]	39.3
a_2b_2	−0.0515	[−0.0874, −0.0213]	14.4
a_3b_3	−0.0624	[−0.0989, −0.0320]	17.4
a_4b_4	−0.0273	[−0.0500, −0.0089]	7.6
Males (n = 575)			
Total effect (c)	−0.3825	[−0.5006, −0.2643]	100
Direct effect (c′)	−0.0498	[−0.1604, 0.0608]^{NS}	13.0
Total indirect effect	−0.3326	[−0.4256, −0.2462]	87.0
a1b1	−0.1809	[−0.2490, −0.1178]	47.3
a2b2	−0.0505	[−0.0874, −0.0203]	13.2
a3b3	−0.0855	[−0.1322, −0.0469]	22.4
a4b4	−0.0158	[−0.0344, −0.0005]	4.1

Abbreviations: CI, confidence interval. NS, not statistically significant effect at the 5% level of significance. Note: Estimates of direct effect, indirect effect, total effect, and proportion of effect mediated were obtained from quadruple mediation analysis in series.
Total indirect effect was the sum of all indirect paths a1b1, a2b2, a3b3, and a4b4.

4.3 STUDY IV

In 2011, 7765 workers participated in the Work Environment Survey, with 3896 reporting musculoskeletal symptoms, of whom 144 (3.7%) received disability pension during the subsequent year. In 2013, 4774 workers responded to the Work Environment Survey, with 2406 reporting musculoskeletal symptoms and 77 (3.2%) receiving disability pension the following year. In all scenarios, data from 2011 served as the training set, while 2013 served as the validation (test) dataset.

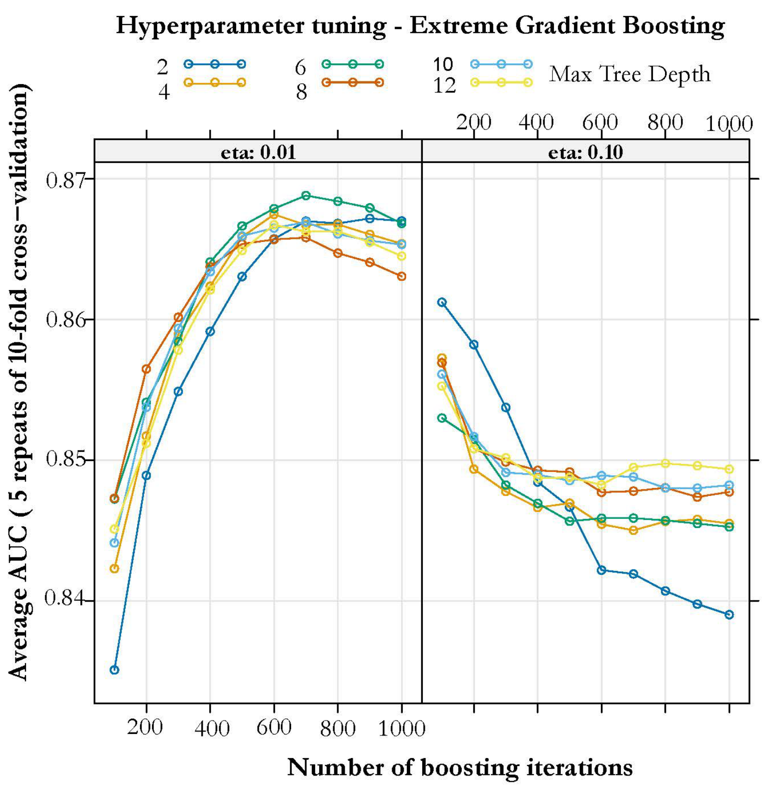

Figure 4.4:The result from tuning hyperparameters for Extreme Gradient Boosting. Max tree depth is the maximum depth of each tree, eta is the learning rate, and the number of boosting iterations determines how many trees are sequentially added to the ensemble.

All machine learning techniques were tuned to optimize the performance of the techniques. Figure 4.4 provides an illustrative example of tuning the

Extreme Gradient Boosting algorithm through the manipulation of three distinct hyperparameters. The x-axis represents the number of boosting iterations, while the different lines with varying colors depict the maximum tree depth. Additionally, the hyperparameter “eta” (learning rate) is represented by two different values, namely 0.01 and 0.10. The average Area Under the Curve (AUC), obtained from a set of candidate values using a 5-repeats of 10-fold cross-validation, is presented. For the given example, a small learning rate (eta = 0.01) with a tree depth of 6 (green line) with 700 boosting iterations resulted in the largest mean AUC of almost 0.87 (left pane of Figure 4.4).

Across all models, ML techniques demonstrated outstanding performance on the training set with c-statistic (AUC) ≥ 0.82 for all models. While there was a slight decline in discrimination performance for the validation set compared to the training set, all techniques maintained similar predictive performance (c-statistics, AUC ≥ 0.85), except for random forest and support vector machines, which exhibited comparatively lower performance (Table 4.8). With an increased number of predictors, a slight enhancement in performance was observed across all ML methods, with AUC values ranging from 0.80 to 0.89 (Model 3). The predicted probability distributions between logistic regression and ML techniques were similar.

The overall performance of ML models, particularly AdaBoost, Gradient Boosting Machine, and Extreme Gradient Boosting, exhibited slight improvements compared to other models, reflected in c-statistics (Table 4.8).

Table 4.8: **Predictive discrimination: c-statistic** (95% CI) derived from machine learning techniques across differing numbers of predictors to forecast the occurrence of disability pension.

Technique	Model 1	Model 2	Model 3
Logistic regression	0.85 (0.80 – 0.90)	0.87 (0.83 – 0.92)	0.87 (0.83 – 0.91)
Random Forest	0.77 (0.71 – 0.83)	0.80 (0.74 – 0.86)	0.81 (0.76 – 0.87)
AdaBoost	0.85 (0.80 – 0.90)	0.87 (0.82 – 0.92)	0.89 (0.85 – 0.93)
Gradient Boosting Machine	0.85 (0.81 – 0.90)	0.86 (0.81 – 0.91)	0.87 (0.83 – 0.91)
Extreme Gradient Boosting	0.86 (0.81 – 0.90)	0.87 (0.83 – 0.92)	0.88 (0.84 – 0.92)
Support Vector Machines	0.67 (0.59 – 0.75)	0.76 (0.69 – 0.83)	0.80 (0.75 – 0.85)
Naïve Bayes Classifiers	0.86 (0.81 – 0.90)	0.87 (0.82 – 0.92)	0.86 (0.82 – 0.90)

Model 1: Seven expert-selected variables.
Model 2: Thirteen variables, demographics, and occupation-related factors
Model 3: Thirty-two variables were selected using the Boruta method, representing work environments and demographics.

Calibration analysis, evaluated by scaled Brier scores (Table 4.9), showed that logistic regression, Naïve Bayes, and Random Forest produced well-calibrated models. AdaBoost achieved marginally better scaled Brier scores than the rest having negative scaled Brier Score for some scenarios. The negative scaled

Table 4.9: **Calibration: Scaled Brier score** derived from machine learning techniques across differing numbers of predictors to forecast the occurrence of disability pension. Lower scores are better.

Technique	Model 1	Model 2	Model 3
Logistic regression	5%	5%	6%
Random Forest	6%	3%	9%
AdaBoost	13%	16%	14%
Gradient Boosting Machine	12%	-2%	8%
Extreme Gradient Boosting	13%	5%	-8%
Support Vector Machines	1%	-15%	-22%
Naïve Bayes Classifiers	5%	2%	2%

Model 1: Seven expert-selected variables.
Model 2: Thirteen variables, demographics, and occupation-related factors
Model 3: Thirty-two variables were selected using the Boruta method, representing work environments and demographics.

Brier scores indicate poorer performance than the base model (proportion of disability pension in the validation dataset).

Calibration analysis, assessed by Spiegelhalter's z-test revealed that logistic regression and Naïve Bayes produced well-calibrated models ($p > 0.05$) across most scenarios (Table 4.10). It means, their predictions were more reliable and exhibited a stronger agreement with the observed outcomes compared to the other machine learning methods. Conversely, AdaBoost, Gradient boosting machine, and Extreme Gradient Boosting exhibited the highest predictive performance. At the same time, Support Vector Machines displayed the lowest predictive performance and calibration, with some negative scaled Brier scores and significant Spiegelhalter's z-test results.

Table 4.10: **Calibration: Spiegelhalter's z-test** from machine learning techniques across differing numbers of predictors to forecast the occurrence of disability pension. A significant result suggests an improperly calibrated model.

Technique	Model 1	Model 2	Model 3
Logistic regression	0.05 (0.96)	0.14 (0.88)	0.64 (0.52)
Random Forest	1.25 (0.21)	2.67 (0.01)	2.44 (0.01)
AdaBoost	-0.20 (0.84)	0.79 (0.43)	2.13 (0.03)
Gradient Boosting Machine	0.25 (0.80)	5.82 (6E-09)	3.9 (9.8E-05)
Extreme Gradient Boosting	0.29 (0.77)	3.75 (2E-04)	20.02 (3.9E-89)
Support Vector Machines	1.08 (0.28)	6.67 (2E-11)	13.39 (6.9E-41)
Naïve Bayes Classifiers	-0.29 (0.77)	-0.95 (0.34)	1.05 (0.3)

Model 1: Seven expert-selected variables.

Model 2: Thirteen variables. demographics. and occupation-related factors

Model 3: Thirty-two variables were selected using the Boruta method. representing work environments and demographics.

Incorporating demographic data and additional predictors enhanced predictive performance but led to poorer calibration, evidenced by negative scaled Brier scores and significant Spiegelhalter's z-test results. While ML techniques demonstrated excellent discrimination, there was limited agreement between observed outcomes and predictions, as illustrated in the calibration statistics (Scaled Brier score and Spiegelhalter's z-test).

5 DISCUSSION

5.1 METHODOLOGICAL CONSIDERATION

The four prospective cohort studies presented in this thesis each pose distinct strengths and limitations. This section will examine the methodological strengths and limitations of the validity and precision of these studies.

The primary objective of an epidemiological study is to produce an effect measure that is both valid and precise [109]. This measure should be generalizable across relevant target populations and consider potential factors that may modify or falsely influence it. The theoretical ideal is to design a flawless study, free from errors, encompassing all individuals in the target population and conducted over an optimal duration to yield an accurate effect measure. However, these ideals are often unattainable in real-life studies due to various limitations.

The accuracy of an effect measure is contingent upon minimizing errors in effect assessment. Drawing inspiration from the hierarchical step-model for causation of bias [110], one can conceptualize a real-life study as gradually deviating from the theoretical ideal as it progresses towards determining the calculated or adjusted effect measure [110]. This model outlines four main steps toward attaining an adjusted effect measure.

5.1.1 CONFOUNDING

A confounding variable is associated with both the exposure and the outcome. If not addressed, confounding can lead to an overestimation or underestimation of an effect, a reversal of the direction of an effect, or the indication of a nonexistent connection. When a variable is strongly linked to exposure but has a weaker association with the outcome of interest, controlling for it can result in over-adjustment and a decrease in the amount of information retained during the analysis.

Since gynecological cancer exclusively affects women, gender did not need to be considered as a potential confounder in Studies I and II. Study I included age as a cofounder, while Study II incorporated a causal model that accounted for four relevant confounders. In Study II, the potential impact of unmeasured confounding was estimated using mediational E-values. Due to the unavailability of data on occupation and the work environment, adjustments for these factors could not be made during the analysis. Therefore, it is crucial to consider the potential influence of unmeasured confounding when interpreting the findings of these studies.

In Study III, the findings may have limited generalizability due to the lower response rate among men. However, conducting separate analyses for men and women provides insights into the influence of gender on the results. Failure to consider potential confounders in the theoretical framework can introduce biases into the estimates. Confounding was not a concern for the comparison of the performance of machine learning methods in Study IV.

5.1.2 MISREPRESENTATION

The subsequent step in the hierarchical model pertains to misrepresentation, where biases may arise due to non-participation and selection bias. It is worth noting that Study I, II, and IV had high participation rates, thus reducing the risk of biases caused by selection.

One notable strength of Study I and II is the inclusion of a large population-based patient cohort from two prominent hospitals in Sweden, which serves a catchment population of 3.5 million individuals. In Sweden's government-funded universal healthcare system, women diagnosed with gynecological cancer are consistently referred to a designated radiotherapy hospital based on their residential location. It is important to examine whether any systematic dropouts during the recruitment process could potentially impact the validity of our findings. Furthermore, the use of official records on disability pensions eliminates attrition bias.

Many cancer survivors did not fulfill the specified eligibility criteria and were consequently omitted from the study cohort. Pertinent data regarding the well-being, health status, and causes of mortality of these excluded women remains unknown, leaving us to merely speculate that the elderly women, as well as those experiencing cancer recurrence or who died before the follow-up period, potentially belonged to a less healthy subgroup. Should the survivors who responded to the questionnaire exhibit superior health and quality of life in comparison to their deceased counterparts who were unable to participate, it is conceivable that the observed effects may have been underestimated.

In Study III, participants were randomly selected from the population to reduce selection bias. Importantly, selection bias related to employment status was minimized by inviting unemployed individuals to participate in the cohort. However, their non-response to the work ability question led to their exclusion from the analyses. It is noteworthy that a lower response rate was observed among men, which could potentially limit the generalizability of the findings due to participation bias. Moreover, an additional analysis comparing individuals who dropped out between baseline and follow-up with those who continued participation revealed no significant differences, suggesting a low risk of attrition bias.

In the context of comparing the performance of machine learning methods in Study IV, the issue of misrepresentation becomes irrelevant.

5.1.3 MISCLASSIFICATION

According to the hierarchical model, the potential for bias increases when the collected information is flawed or misclassified. This phenomenon, also known as information bias or measurement bias, is particularly relevant in studies that rely on self-reported data, where the questionnaire plays a crucial role.

Non-differential misclassification of exposure occurs when the misclassification of the exposure is unrelated to the outcome of interest. This type of misclassification tends to introduce a bias towards the null in the estimate. On the other hand, differential misclassification arises when exposure misclassification varies between participants with and without the outcome. This type of misclassification leads to a biased estimate that could either lean toward or diverge away from the null hypothesis, conditional on the percentages of participants misclassified.

Studies I-II and IV utilized disability pension records from official registers. The remaining self-reported data was collected using validated questions [111-113]. The involvement of cancer survivors in the questionnaire development stage and the extensive face validation within the study population suggest a high likelihood of respondents understanding and answering the questions as intended, thereby reducing the risk of misclassification in Studies I and II.

One limitation of Study III is the lack of validation for certain variables. In Study IV, musculoskeletal symptoms and work demands were self-reported, and a review indicated varying levels of validity among survey items. Some items demonstrated robust validity, while others showed lower validity. In the context of comparing the performance of machine learning methods, the phenomenon of misclassification becomes none relevant.

For all studies, single-item questions may provide valuable insights, although they may not capture complex constructs as comprehensively as multi-question scales. However, these questions were designed to closely approximate the represented dimensions to the best extent possible. Responding to a postal questionnaire in ample time, independently and in privacy, minimizes potential biases induced by interviewer influence or external pressures.

5.1.4 ANALYTICAL ADJUSTMENT

The last step entails making analytical adjustments to minimize the impact of confounding factors and biases. Wherever applicable, statistical adjustments were used to align the groups under study, thus improving the comparability of individuals who were exposed with those who were not in the cohort.

5.2 RANDOM ERRORS

Random errors, which arise from the inherent variability in the data, are unpredictable and have an inverse relationship with the sample size. These errors have the potential to mask true associations and affect the accuracy of effect estimates. The evaluation of accuracy is performed using confidence intervals, where a wider interval indicates lower accuracy, while a narrower interval suggests higher accuracy. In our research studies, the 95% confidence intervals accurately reflect the variability observed in our point estimates. Nevertheless, the precision of these estimates is influenced by the sample size, and in specific analyses (such as Study II), the limited number of individuals per category resulted in broader confidence intervals for some effect measures.

5.3 ETHICAL CONSIDERATIONS

This research was conducted following the Declaration of Helsinki [114]. All studies received approval from the Regional Ethical Review Board in Gothenburg, Sweden

- Study I & II with "Registration number 691-17"
- Study III with "Registration number T876-11 and T862-17".
- Study IV with "Registration number 221–15".

Despite the observational nature of these studies, ethical scrutiny remains crucial. As researchers, we are obligated to consider the impact of our research on subjects' safety, integrity, and dignity. The studies in this thesis did not expose participants to physical interventions, thereby ensuring their safety. All data was anonymized, and results were reported at an aggregated level.

Researchers must maintain objectivity, transparency, and honesty when reporting study results. All significant aspects of the study should be reported and clarified. In our studies, we adhered to established guidelines such as SAMPL [115], STROBE [116], and AGReMA Checklist [117] to ensure comprehensive and transparent reporting of results.

The decline in work ability has profound implications for the worker, their families, and society at large. Given this context, it can be concluded that the benefits of this research will have a substantial impact.

5.4 FINDINGS AND INTERPRETATIONS

Study I and II

To the best of our knowledge, Studies I and II were the first studies that examined the relationship between radiation-induced gastrointestinal syndromes and disability pension within a national cohort. The findings of these studies are of great significance due to the possibility of modifying radiation-induced syndromes. These syndromes can be predicted based on the doses of ionizing radiation received by different parts of the intestine. Through our analysis of Swedish population-based registers and patient-reported outcomes, we discovered a significant association between specific radiation-induced syndromes and disability pension. It is worth noting that survivors of gynecological cancer with these syndromes were twice as likely to receive a disability pension compared to those without, which is a novel finding in this field. While previous research has suggested an increased probability of disability pension among cancer survivors [53, 118, 119], our studies provide detailed insights into the risks associated with specific syndromes, particularly defecation urgency, fecal leakage, and anal blood discharge.

The existing literature provides additional context, suggesting that radiation-induced urgency syndrome and fecal leakage syndrome may have a negative impact on work ability, which is in line with our findings. A previous study has documented that cancer survivors who have undergone radiation therapy experience diarrhea and bowel symptoms [120, 121], increased work absenteeism, and reduced physical capabilities among cancer survivors that undergone radiation therapy [122]. Gastrointestinal problems have been shown to have a negative effect on social functioning [18, 121], quality of life [13, 123], and ability to work [71], highlighting the importance of comprehensively addressing gastrointestinal health among gynecological cancer survivors.

Furthermore, our study reveals a dose-response relationship between the number of radiation-induced syndromes and the likelihood of receiving a disability pension. This finding aligns with earlier observations that having multiple treatment-related syndromes exacerbates work ability limitations among cancer survivors [124]. The collective effect of these syndromes emphasizes their significant influence on survivors' employment and work ability, thereby affecting their participation in the workforce.

In a related discourse, the integration of self-reported and register data has unveiled the mediating role of various aspects of quality of life in the relationship between radiation-induced urgency syndrome and disability pension. This discovery underscores the importance of quality of life in navigating work-life dynamics among survivors of gynecological cancer. The predominance of global and physical aspects of quality of life in mediating this relationship highlights their crucial role, as urgency syndrome exerting a noticeable negative impact. Interestingly, psychological aspects play a lesser mediating role, which is consistent with prior research that emphasizes the priority of physical components of quality of life in predicting disability pension [125].

These findings corroborate the broader discourse on the interplay between syndrome-specific symptoms, quality of life, and work ability among gynecological cancer survivors. By shedding light on the mediating mechanisms underlying the relationship between urgency syndrome and the need for disability pension, our study contributes to a nuanced understanding of survivorship dynamics and employment outcomes. Recognizing the diverse impact of syndromes resulting from radiation and implementing interventions that focus on improving quality of life could potentially reduce the likelihood of disability pension and improve the work-life trajectory of cancer survivors.

Study III

Our examination of baseline self-perceived stress and its impact on future work ability among young adults reveals intriguing insights. Our findings emphasize the mediating roles of ongoing self-perceived stress, work-related demands encroaching on personal life, feelings of control over personal life, and feeling well-rested upon waking. Collectively, these factors explain 81% of the observed association between stress and work ability, highlighting the crucial role of achieving a harmonious equilibrium between personal and work-life domains in alleviating the detrimental consequences of stress on work ability.

In comparison, only a few studies have explored the mediating dimensions of work ability [126, 127]. Although previous research has acknowledged the associations between stress and work ability [19-22], there has been an insufficient exploration of potentially related variables that mediate this relationship. Our study bridges this gap by revealing a significant, moderate correlation between stress and work ability over five years, aligning with previous observations suggesting the long-term impacts of stress on work ability trends. Our findings clarify the role of selected mediating factors in the quadruple mediation model, explaining their contribution to this association.

Our investigation goes beyond traditional views on work-home interference by examining the influence of work demands and perceived control over private life on the stress-work ability relationship. While previous studies offer varied perspectives on the stress-work-home interface [128, 129], the precise dynamics remain to be discovered. Discrepancies persist regarding the directionality of stress and work-home interference [128], underscoring the need for further investigations. Our focus on young workers emphasizes the importance of mitigating work spillover into private life and preserving feelings of control.

In summary, our findings highlight the multifaceted nature of the stress-work ability relationship. They emphasize the importance of addressing work-life balance dynamics and enhancing feelings of control to optimize work ability among young adults.

Study IV

Study IV provides valuable insights into the prediction accuracy of logistic regression and other machine learning techniques in predicting disability pensions among Swedish workers with musculoskeletal symptoms. Although the study did not focus on predicting disability pension, the importance of selected variables in predicting disability pension is a significant secondary finding. The use of data from a representative sample of the Swedish labor market, combined with official registers, increases the robustness of the prediction models. The credibility of our prediction model is further reinforced by subjecting the models to temporal external validation.

Our primary observation underscores the comparability of logistic regression with more advanced machine-learning techniques across all models. All machine learning techniques, apart from Random Forest and Support Vector Machines demonstrated high discriminative performance (AUC $\geq$ 0.85). This suggests a strong likelihood (85%) that the machine learning technique will rank a worker receiving a disability pension higher than one without. Boosting algorithms, such as AdaBoost, Gradient Boosting Machine, and Extreme Gradient Boosting, showed superior performance, possibly because of their capability to capture complex interactions and learning from the mistakes of previous models sequentially. Despite not considering interactions, logistic regression exhibited comparable performance, indicating that the overall importance of interactions in this context is limited.

Our findings are consistent with previous studies, underscoring the comparable discriminative ability of logistic regression to state-of-the-art machine learning techniques [130, 131]. Despite debates about the trade-off between discrimination and calibration [132], our study reaffirms the practicality of logistic regression in prediction modeling. It is important to recognize that low values of the Brier score do not necessarily imply better calibration. While a low Brier score indicates improved accuracy in predictions, it does not guarantee optimal calibration [133]. Calibration assessment requires additional inspection, e.g., Spiegelhalter's z-test [133].

While our study assesses commonly used machine learning techniques, the rapidly advancing field of technology introduces additional methods that could potentially yield different results. The effectiveness of each method may vary depending on the characteristics of the dataset, which necessitates caution when generalizing results. Modern machine learning methods might offer limited contributions in biomedical studies where the signal-to-noise ratio could be lower [130], i.e. the noise is relatively stronger compared to the signal.

Moreover, our results were based on hyperparameter tuning with specific candidate values. This leaves room for enhancing predictive ability through more refined hyperparameter tuning. Despite these limitations, our study makes a significant contribution to the emerging field of machine learning in occupational health.

6 CONCLUSION

There exists a complex relationship between health and work ability. Radiation-induced gastrointestinal syndrome influences self-reported quality of life and disability pension, where quality of life mediates part of the gastrointestinal syndrome's effect on disability pension. Occupational factors mediate the impact of perceived stress on work ability in young adults. Also, feelings of control over personal life are important for maintaining work ability in young adults with stress complaints. The performance of multivariable logistic regression, which utilizes clinically relevant predictors, is comparable to that of sophisticated machine learning algorithms when it comes to predicting disability pensions.

7 FUTURE PERSPECTIVES

The findings from the four studies presented in this thesis offer valuable insights into the potential associations between the investigated exposures, outcomes, and mediators. The observed frequency of disability pension and the worsened quality of life among gynecological cancer survivors reporting radiation-induced gastrointestinal syndromes highlight the need for rigorous investigations using diverse research methodologies. The current body of literature on aspects of occupational life and quality of life following radiation therapy for pelvic cancer is limited, particularly in studies that combine survivors' perspectives with quantitative data.

Identifying mediation pathways linking urgency syndrome and disability pension emphasizes the importance of addressing dimensions of quality of life, such as physical health and psychological well-being, in survivorship care programs. Future research should explore comprehensive interventions targeting these mediators to mitigate the risk of disability pension among gynecological cancer survivors. Further investigation into the mechanisms underlying the association between specific syndromes, quality of life, and disability pension receipt could shed light on potential pathways for intervention and support services tailored to the needs of survivors. Treatment plans should incorporate strategies for managing working life and aspects of quality of life after treatment.

The longitudinal relationship between stress and work ability among workers underscores the significance of psychosocial work environment factors in occupational health. Future research should prioritize testing interventions targeting mediating factors to confirm their effectiveness in preventing declines in work ability over time. Interventions aimed at reducing long-term stress and promoting a supportive work environment could mitigate the negative impact of stress on work ability.

The predictive performance of machine learning techniques in identifying workers at risk of disability pension highlights the potential of machine learning approaches in occupational health surveillance. Another aspect is the need for reliable information about the work environment, including health, functional capacities, competence, values, attitudes, and motivation. Additionally, information on factors outside of work, such as family, close community, other societal environments, services, and legislation, is necessary to promote sustainable work ability. Future studies could explore the integration of such information to enhance the accuracy and calibration of predictive models, thereby facilitating early identification and intervention for workers with musculoskeletal symptoms at risk of work disability.

8 ACKNOWLEDGMENT

The studies in this thesis were performed at the School of Public Health and Community Medicine, University of Gothenburg, Gothenburg, Sweden.

One of the most rewarding parts of working on my thesis was the opportunity to meet and collaborate with a nice and fascinating group of individuals. Their contributions not only enriched the content of my thesis but also made the journey delightful. I thank you all and extend my heartfelt gratitude to:

Helena Sandén, my main supervisor, for providing knowledgeable guidance, engaging in stimulating discussions, offering constructive advice, and demonstrating meticulous attention to detail. Your untiring support has been an endless source of encouragement.

Martin Adiels, my co-supervisor, for introducing me to the fascinating world of statistical learning. Our discussions have been intellectually stimulating, and your constructive ideas have enriched my professional development. Your enthusiasm for my growth has been invaluable.

Mats Hagberg, my co-supervisor, for introducing me to the intriguing field of occupational health, engaging conversations, and patience. Your inspiration and encouragement were instrumental in my decision to pursue this Ph.D. education. I am deeply grateful for the opportunities you made possible.

To all three of you, thank you for your unwavering trust and support throughout this journey!

Esteemed co-authors, **Sara Thomée**, **Gunnar Steineck**, **Art van Schaaijk**, **Monique Frings-Dresen**, and **Karen Nieuwenhuijsen.** Their expertise, invaluable contributions, and positive collaborations have enriched this work.

Max Petzold for being a great mentor and inspiration. **Mikael Svensson** for generous support. **Maria Åberg** for fostering a positive work environment and being a fantastic manager. **Catrin Wessman** and **Anna Rehammar** for flexibility in duty planning, which allowed me to focus on my Ph.D. studies.

Eva Andersson and **Leo Stockfelt** for being opponents during the mock defense (*predisputation*) and for thought-provoking questions and insightful suggestions. Eva Andersson for much appreciated constructive feedback on the thesis manuscript. **Ying Li, Barbora Kessel**, and **Qinyun Lin** for beneficial feedback on my studies. **Jonas Björk** and **Gunnel Hensing** for being a part of the assessment committee during my half-time seminar and for suggestions for improvements.

Anna Grimby-Ekman for being a friend, sharing laughs, frustrations, and adventures related to my work, and making this journey memorable. **Ann-Sofie Liljenskog-Hill, Cecilia Groglopo**, and **Fredrik Petersson**, for your humor, friendship, and assistance, and for celebrating the successes. **Kristina Aurelius** (formerly **Gyllensten**) and **Hatice Koca** for enriching conversations that extended beyond work. **Laith Hussain** and **Stefan Oliv** for engaging collegially and providing valuable insights about the Ph.D. journey.

Everyone in the research group "**Work Ability and Ergonomics**", "**Biostatistics**", and the "**ViBIT/SUNDS**" network, as well as my supportive **colleagues** and **affiliates** at **Occupational and Environmental Medicine** (**AMM**), and the **School of Public Health and Community Medicine.** No one mentioned, no one forgotten. Your camaraderie, stimulating intellectual exchange, and encouragement have fostered a nurturing and supportive environment—one that promotes scholarly growth. And to all the wonderful individuals at SND, with a special mention to **Elisabeth Strandhagen** and **Lisa Isaksson,** for the enriching conversations and support.

Fellow **Ph.D. students** from **SPHCM**, and those I've had the pleasure of meeting through courses or while working in the PhD Student Council at the Sahlgrenska Academy and the Doctoral Committee at the Institute of Medicine, I have enjoyed our discussions, and the good company has made this journey remarkable!

My **dear friends** and **relatives**, you know who you are. As the thesis journey reaches its completion, I eagerly anticipate sharing more moments with each of you. Let's embrace and celebrate this milestone together!

My dear parents, **Mehr-un-Nisa** and **Noor Muhammad**, my unwavering supporters, and true believers. You have been a constant source of motivation

and your love, a comforting presence in times of doubt. To my **brother** and **sister**, your love and presence have meant the world to me, and our shared moments have been filled with joy.

My beloved **family** for their understanding and encouragement, which have sustained me throughout this journey. For providing me the time and opportunity to face the challenges of my work and for sharing with me both the hardships and joy at every stage of this project.

To my dear wife, "***Du är min finaste stund***"!

9 REFERENCES

1. Lundin, A., et al., *The Association Between Self-Assessed Future Work Ability and Long-Term Sickness Absence, Disability Pension and Unemployment in a General Working Population: A 7-Year Follow-Up Study.* Journal of Occupational Rehabilitation, 2016. **26**(2): p. 195-203.
2. Jääskeläinen, A., et al., *Work ability index and perceived work ability as predictors of disability pension: a prospective study among Finnish municipal employees.* Scandinavian Journal of Work, Environment & Health, 2016. **42**(6): p. 490-499.
3. Alavinia, S.M., et al., *Determinants of work ability and its predictive value for disability.* Occupational Medicine, 2008. **59**(1): p. 32-37.
4. Gould, R., J. Ilmarinen, J. Järvisalo, and S. Koskinen, *Dimensions of Work Ability : Results of the Health 2000 Survey*. Dimensions of Work Ability. 2008, Helsinki.
5. van den Berg, S., A. Burdorf, and S.J.W. Robroek, *Associations between common diseases and work ability and sick leave among health care workers.* International Archives of Occupational and Environmental Health, 2017. **90**(7): p. 685-693.
6. Rice, N.E., I.A. Lang, W. Henley, and D. Melzer, *Common health predictors of early retirement: findings from the English Longitudinal Study of Ageing.* Age and Ageing, 2010. **40**(1): p. 54-61.
7. de Wind, A., et al., *Pathways through which health influences early retirement: a qualitative study.* BMC Public Health, 2013. **13**(1): p. 292.
8. Karpansalo, M., et al., *Perceived health as a predictor of early retirement.* Scandinavian Journal of Work, Environment & Health, 2004(4): p. 287-292.
9. Bevan, S., *Economic impact of musculoskeletal disorders (MSDs) on work in Europe.* Best Practice & Research Clinical Rheumatology, 2015. **29**(3): p. 356-373.
10. Steineck, G., et al., *Toward Restored Bowel Health in Rectal Cancer Survivors.* Seminars in Radiation Oncology, 2016. **26**(3): p. 236-250.
11. Andreyev, H.J.N., *Gastrointestinal problems after pelvic radiotherapy: the past, the present and the future.* Clinical Oncology, 2007. **19**(10): p. 790-9.
12. Pfaendler, K.S., L. Wenzel, M.B. Mechanic, and K.R. Penner, *Cervical Cancer Survivorship: Long-term Quality of Life and Social Support.* Clinical Therapeutics, 2015. **37**(1): p. 39-48.
13. Abayomi, J., J. Kirwan, and A. Hackett, *The prevalence of chronic radiation enteritis following radiotherapy for cervical or endometrial*

cancer and its impact on quality of life. European Journal of Oncology Nursing, 2009. **13**(4): p. 262-267.

14. Bergmark, K., et al., *Patient-rating of distressful symptoms after treatment for early cervical cancer.* Acta Obstetricia et Gynecologica Scandinavica, 2002. **81**(5): p. 443-443.
15. Baloch, A.N., et al., *Disability pension among gynaecological cancer survivors with or without radiation-induced survivorship syndromes.* Journal of Cancer Survivorship, 2021. **16**(4): p. 834-843.
16. de Boer, A.G., et al., *Work ability and return-to-work in cancer patients.* British Journal of Cancer, 2008. **98**(8): p. 1342-7.
17. Hauer-Jensen, M., J.W. Denham, and H.J.N. Andreyev, *Radiation enteropathy--pathogenesis, treatment and prevention.* Nature Reviews Gastroenterology & Hepatology 2014. **11**(8): p. 470-479.
18. Dunberger, G., et al., *Fecal incontinence affecting quality of life and social functioning among long-term gynecological cancer survivors.* International Journal of Gynecologic Cancer, 2010. **20**(3): p. 449-460.
19. Yang, T., et al., *Effect of Stress on the Work Ability of Aging American Workers: Mediating Effects of Health.* International Journal of Environmental Research and Public Health, 2019. **16**(13): p. 2273.
20. Marklund, S., et al., *Work ability and productivity among dentists: associations with musculoskeletal pain, stress, and sleep.* International Archives of Occupational and Environmental Health, 2020. **93**(2): p. 271-278.
21. Bernburg, M., K. Vitzthum, D.A. Groneberg, and S. Mache, *Physicians' occupational stress, depressive symptoms and work ability in relation to their working environment: a cross-sectional study of differences among medical residents with various specialties working in German hospitals.* BMJ Open, 2016. **6**(6): p. e011369.
22. Lindegård, A., P. Larsman, E. Hadzibajramovic, and G. Ahlborg, *The influence of perceived stress and musculoskeletal pain on work performance and work ability in Swedish health care workers.* International Archives of Occupational and Environmental Health, 2014. **87**(4): p. 373-379.
23. Leclerc, C., F. De Keulenaer, and S. Belli, *OSH Pulse Occupational safety and health in post-pandemic workplaces*, in *Flash Eurobarometer*. 2022, European Agency for Safety and Health at Work. p. 55.
24. Hill, E.J., et al., *Work and Family over the Life Course: Do Older Workers Differ?* Journal of Family and Economic Issues, 2014. **35**(1): p. 1-13.
25. Salminen, S., et al., *Is a single item stress measure independently associated with subsequent severe injury: a prospective cohort study*

of 16,385 forest industry employees. BMC Public Health, 2014. **14**(1): p. 543.

26. Veldman, K., et al., *A life course perspective on mental health problems, employment, and work outcomes.* Scand J Work Environ Health, 2017. **43**(4): p. 316-325.
27. Melchior, M., et al., *Work stress precipitates depression and anxiety in young, working women and men.* Psychological Medicine, 2007. **37**(8): p. 1119-1129.
28. Jackson, J.A., et al., *Risk factors for surgically treated cervical spondylosis in male construction workers: a 20-year prospective study.* The Spine Journal, 2023. **23**(1): p. 136-145.
29. Karampampa, K., et al., *Healthcare, Sickness Absence, and Disability Pension Cost Trajectories in the First 5 Years After Diagnosis with Multiple Sclerosis: A Prospective Register-Based Cohort Study in Sweden.* PharmacoEconomics - Open, 2020. **4**(1): p. 91-103.
30. Bradley, C.J., L. Penberthy, K.J. Devers, and D.J. Holden, *Health Services Research and Data Linkages: Issues, Methods, and Directions for the Future.* Health Services Research, 2010. **45**(5p2): p. 1468-1488.
31. Waddell, G. and A.K. Burton, *Is work good for your health and well-being?*, ed. A.K. Burton. 2006, London: London : TSO (The Stationery Office).
32. Centralbyrån, A.S., *Arbetsorsakade besvär 2020 [Work-Related Disorders 2020]*, in *Arbetsmiljöstatistik* A.P. Klevestedt, Editor. 2021, Arbetsmiljöverket; Statistiska Centralbyrån: Stockholm, Sweden.
33. Ludvigsson, M., T. Svensson, and K. Alexanderson, *Begreppet arbetsförmåga : en litteraturgenomgång*. 2006, Stockholm: Arbetslivsinstitutet.
34. Ilmarinen, J., K. Tuomi, and J. Seitsamo, *New dimensions of work ability.* International Congress Series, 2005. **1280**: p. 3-7.
35. Agency, T.S.I. *Sickness benefit*. 2023 2024-01-16]; Available from: https://www.forsakringskassan.se/english/sick/employee/sickness-benefit.
36. Nordenfelt, L., *The concept of work ability*. 2008: PIE Peter Lang. 234.
37. Tengland, P.-A., *The Concept of Work Ability.* Journal of Occupational Rehabilitation, 2011. **21**(2): p. 275-285.
38. Ilmarinen, J., *From Work Ability Research to Implementation.* International Journal of Environmental Research and Public Health, 2019. **16**(16): p. 2882.
39. Ilmarinen, J., *The Work Ability Index (WAI).* Occupational Medicine, 2007. **57**(2): p. 160-160.
40. Bethge, M., K. Spanier, S. Köhn, and A. Schlumbohm, *Self-reported work ability predicts health-related exit and absence from work, work*

participation, and death: longitudinal findings from a sample of German employees. International Archives of Occupational and Environmental Health, 2021. **94**(4): p. 591-599.
41. Lundin, A., et al., *Predictive validity of the Work Ability Index and its individual items in the general population.* Scandinavian Journal of Public Health, 2017. **45**(4): p. 350-356.
42. Martus, P., et al., *A comparative analysis of the Work Ability Index.* Occupational Medicine, 2010. **60**(7): p. 517-524.
43. Ahlstrom, L., et al., *The work ability index and single-item question: associations with sick leave, symptoms, and health – a prospective study of women on long-term sick leave.* 2010(5): p. 404-412.
44. Kinnunen, U. and J. Nätti, *Work ability score and future work ability as predictors of register-based disability pension and long-term sickness absence: A three-year follow-up study.* Scandinavian Journal of Public Health, 2018. **46**(3): p. 321-330.
45. Karlsson, N., et al., *Risk of disability pension in relation to gender and age in a Swedish county; a 12-year population based, prospective cohort study.* Work, 2006. **27**: p. 173-179.
46. Kadefors, R., et al., *Social inequality in working life expectancy in Sweden.* Z Gerontol Geriatr, 2019. **52**(Suppl 1): p. 52-61.
47. Ludvigsson, J.F., et al., *The longitudinal integrated database for health insurance and labour market studies (LISA) and its use in medical research.* European Journal of Epidemiology, 2019. **34**(4): p. 423-437.
48. Agency, T.S.I. *Sick for one year or longer*. 2023 [cited 2024 2024-01-16]; Available from: https://www.forsakringskassan.se/english/moving-to-working-studying-or-newly-arrived-in-sweden/social-insurance-system.
49. Agency, T.S.I. *Activity compensation and sickness compensation*. [Webpage] 2024 [cited 2024 2024-03-09]; Information on disability pension in Sweden]. Available from: https://www.forsakringskassan.se/english/disability/activity-compensation-and-sickness-compensation.
50. Feuerstein, M., et al., *Work in cancer survivors: a model for practice and research.* J Cancer Surviv, 2010. **4**(4): p. 415-37.
51. Mehnert, A., *Employment and work-related issues in cancer survivors.* Critical Reviews in Oncology/Hematology, 2011. **77**(2): p. 109-130.
52. Lind, H., et al., *Late symptoms in long-term gynaecological cancer survivors after radiation therapy: a population-based cohort study.* Br J Cancer, 2011. **105**(6): p. 737-45.
53. Everhov, A.H., et al., *Lost workdays in uterine cervical cancer survivors compared to the general population: impact of treatment and relapse.* Journal of Cancer Survivship, 2016. **10**(3): p. 514-23.

54. Jex, S.M., T.A. Beehr, and C.K. Roberts, *The meaning of occupational stress items to survey respondents.* Journal of Applied Psychology, 1992. **77**(5): p. 623-628.
55. Ursin, H. and H.R. Eriksen, *The cognitive activation theory of stress.* Psychoneuroendocrinology, 2004. **29**(5): p. 567-592.
56. Oliv, S., A. Noor, E. Gustafsson, and M. Hagberg, *A Lower Level of Physically Demanding Work Is Associated with Excellent Work Ability in Men and Women with Neck Pain in Different Age Groups.* Safety and Health at Work, 2017. **8**(4): p. 356-363.
57. Oliv, S., et al., *Important work demands for reducing sickness absence among workers with neck or upper back pain: a prospective cohort study.* BMC Musculoskeletal Disorders, 2019. **20**(1): p. 529.
58. Karpansalo, M., et al., *Physical Workload and Risk of Early Retirement: Prospective Population-Based Study Among Middle-Aged Men.* Journal of Occupational and Environmental Medicine, 2002. **44**(10): p. 930-939.
59. Falkstedt, D., et al., *Disability pensions related to heavy physical workload: a cohort study of middle-aged and older workers in Sweden.* International Archives of Occupational and Environmental Health, 2021. **94**(8): p. 1851-1861.
60. Badarin, K., et al., *Combined exposure to heavy physical workload and low job control and the risk of disability pension: A cohort study of employed men and women in Sweden.* International Archives of Occupational and Environmental Health, 2023. **96**(7): p. 973-984.
61. Lahelma, E., et al., *Working conditions as risk factors for disability retirement: a longitudinal register linkage study.* Bmc Public Health, 2012. **12**.
62. Botey Gaude, L., et al., *Working conditions in the time of COVID-19: Implications for the future*, in *European Working Conditions Telephone Survey 2021 series*. 2022, European Foundation for the Improvement of Living and Working Conditions: Luxembourg.
63. Arbetsmiljöverket; Statistiska Centralbyrån, *Arbetsmiljön 2013 [The Work Environment 2013]*. 2014: Stockholm. p. 456.
64. Bowling, A., *Research Methods in Health: Investigating Health and Health Services*. Fifth edition ed. 2023, Maidenhead: Maidenhead: McGraw-Hill Education.
65. Lenderink, A.F., et al., *Review on the validity of self-report to assess work-related diseases.* Int Arch Occup Environ Health, 2012. **85**(3): p. 229-51.
66. W. Steyerberg, E., *Clinical prediction models: a practical approach to development, validation, and updating*. Second Edition ed. Statistics for Biology and Health. 2008. XXXIII, 558.

67. Wilson, P.W.F., et al., *Prediction of Coronary Heart Disease Using Risk Factor Categories.* Circulation, 1998. **97**(18): p. 1837-1847.
68. Sidey-Gibbons, J.A.M. and C.J. Sidey-Gibbons, *Machine learning in medicine: a practical introduction.* BMC Medical Research Methodology, 2019. **19**(1): p. 64.
69. Hastie, T., R. Tibshirani, J.H. Friedman, and J.H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*. Vol. 2. 2009: Springer.
70. Chan, V.C.H., et al., *The role of machine learning in the primary prevention of work-related musculoskeletal disorders: A scoping review.* Applied Ergonomics, 2022. **98**: p. 103574.
71. Dunberger, G., et al., *Fecal incontinence affecting quality of life and social functioning among long-term gynecological cancer survivors.* Int J Gynecol Cancer, 2010. **20**(3): p. 449-60.
72. Steineck, G., et al., *Identifying radiation-induced survivorship syndromes affecting bowel health in a cohort of gynecological cancer survivors.* PLOS ONE, 2017. **12**(2): p. e0171461.
73. Dunberger, G., et al., *Self-reported symptoms of faecal incontinence among long-term gynaecological cancer survivors and population-based controls.* European Journal of Cancer, 2010. **46**(3): p. 606-615.
74. Ekman, A., et al., *UngVuxen - Basenkät [Young Adults - baseline questionnaire]*, in *Rapport från Arbets- och miljömedicin*. 2008, Occupational and Environmental Medicine: Gothenburg. p. 242.
75. Palmerud, G. and M. Hagberg, *Arbetsförmåga och fysisk och psykosocial exponering i arbetet bland unga vuxna [Work ability, physical and psychosocial work exposure among young adults]*. 2010, Occupational and Environmental Medicine: Gothenburg; Sweden. p. 22.
76. Dovey-Pearce, G., et al., *Young adults' (16–25 years) suggestions for providing developmentally appropriate diabetes services: a qualitative study.* Health & Social Care in the Community, 2005. **13**(5): p. 409-419.
77. Boström, M., J.K. Sluiter, and M. Hagberg, *Changes in work situation and work ability in young female and male workers. A prospective cohort study.* BMC Public Health, 2012. **12**(1): p. 694.
78. Thomée, S., A. Härenstam, and M. Hagberg, *Mobile phone use and stress, sleep disturbances, and symptoms of depression among young adults - a prospective cohort study.* BMC Public Health, 2011. **11**(1): p. 66.
79. Thomée, S., A. Härenstam, and M. Hagberg, *Computer use and stress, sleep disturbances, and symptoms of depression among young adults – a prospective cohort study.* BMC Psychiatry, 2012. **12**(1): p. 176.

80. Sweden, S., *Labour Force Surveys (LFS)*, in *Official Statistics of Sweden*. 2024: Stockholm, Sweden.
81. van Buuren, S. and K. Groothuis-Oudshoorn, *mice: Multivariate Imputation by Chained Equations in R.* Journal of Statistical Software, 2011. **45**(3): p. 1 - 67.
82. Kursa, M.B. and W.R. Rudnicki, *Feature Selection with the Boruta Package.* Journal of Statistical Software, 2010. **36**(11): p. 1 - 13.
83. Steyerberg, E.W. and F.E. Harrell, Jr., *Prediction models need appropriate internal, internal-external, and external validation.* J Clin Epidemiol, 2016. **69**: p. 245-7.
84. Onelov, E., et al., *Measuring anxiety and depression in the oncology setting using visual-digital scales.* Acta Oncol, 2007. **46**(6): p. 810-6.
85. Skoogh, J., et al., *'A no means no'—measuring depression using a single-item question versus Hospital Anxiety and Depression Scale (HADS-D).* Annals of Oncology, 2010. **21**(9): p. 1905-1909.
86. Elo, A.L., A. Leppanen, and A. Jahkola, *Validity of a single-item measure of stress symptoms.* Scand J Work Environ Health, 2003. **29**(6): p. 444-51.
87. Baron, R.M. and D.A. Kenny, *The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations.* Journal of Personality and Social Psychology, 1986. **51**(6): p. 1173-82.
88. MacKinnon, D.P., et al., *A comparison of methods to test mediation and other intervening variable effects.* Psychological Methods, 2002. **7**(1): p. 83-104.
89. Hayes, A.F., *Introduction to mediation, moderation, and conditional process analysis a regression-based approach*. Methodology in the Social Sciences Series, ed. T.D. Little. 2022, New York: New York : Guilford Publications.
90. Richiardi, L., R. Bellocco, and D. Zugna, *Mediation analysis in epidemiology: methods, interpretation and bias.* International Journal of Epidemiology, 2013. **42**(5): p. 1511-1519.
91. VanderWeele, T.J., *Explanation in causal inference : methods for mediation and interaction*. 2015: New York, NY : Oxford University Press. 706.
92. Valeri, L. and T.J. VanderWeele, *Mediation Analysis Allowing for Exposure–Mediator Interactions and Causal Interpretation: Theoretical Assumptions and Implementation With SAS and SPSS Macros.* Psychological Methods, 2013. **18**(2): p. 137-150.
93. VanderWeele, T.J., *Explanation in causal inference : methods for mediation and interaction*. 2015: New York, NY : Oxford University Press.

94. Smith, L.H. and T.J. VanderWeele, *Mediational E-values: Approximate Sensitivity Analysis for Unmeasured Mediator–Outcome Confounding.* Epidemiology, 2019. **30**(6): p. 835-837.
95. Breiman, L., *Statistical Modeling: The Two Cultures.* Statistical Science, 2001. **16**(3): p. 199-215.
96. Peduzzi, P., et al., *A simulation study of the number of events per variable in logistic regression analysis.* Journal of Clinical Epidemiology, 1996. **49**(12): p. 1373-1379.
97. van der Ploeg, T., P.C. Austin, and E.W. Steyerberg, *Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints.* BMC Medical Research Methodology, 2014. **14**(1): p. 137.
98. Wacholder, S., *Binomial regression in GLIM: Estimating risk ratios and risk differences.* American Journal of Epidemiology, 1986. **123**(1): p. 174-184.
99. Skov, T., J. Deddens, M.R. Petersen, and L. Endahl, *Prevalence proportion ratios: estimation and hypothesis testing.* Int J Epidemiol, 1998. **27**(1): p. 91-5.
100. Hosmer, D.W., *Applied logistic regression*. 3rd ed, ed. S. Lemeshow and R.X. Sturdivant. 2013, Chicester: Chicester : Wiley.
101. Jordan, M.I. and T.M. Mitchell, *Machine learning: Trends, perspectives, and prospects.* Science, 2015. **349**(6245): p. 255-260.
102. Wiemken, T.L. and R.R. Kelley, *Machine Learning in Epidemiology and Health Outcomes Research.* Annual Review of Public Health, 2020. **41**(1): p. 21-36.
103. Kuhn, M. and K. Johnson, *Applied predictive modeling*, ed. K. Johnson. 2013, New York, NY: Springer New York, NY.
104. Breiman, L., *Bagging predictors.* Machine Learning, 1996. **24**(2): p. 123-140.
105. Vapnik, V.N., *Statistical learning theory*. 1998, New York: New York : Wiley.
106. Steyerberg, E.W., et al., *Assessing the performance of prediction models: a framework for traditional and novel measures.* Epidemiology, 2010. **21**(1): p. 128-38.
107. Hayes, A.F., A.K. Montoya, and N.J. Rockwood, *The Analysis of Mechanisms and Their Contingencies: PROCESS versus Structural Equation Modeling.* Australasian Marketing Journal, 2017. **25**(1): p. 76-81.
108. Kuhn, M., *Building Predictive Models in R Using the caret Package.* Journal of Statistical Software, 2008. **28**(5): p. 1 - 26.
109. Rothman, K.J., *Modern epidemiology*. 3. ed, ed. S. Greenland and T.L. Lash. 2008: Philadelphia : Lippincott Williams & Wilkins.

110. Steineck, G., H. Hunt, and J. Adolfsson, *A hierarchical step-model for causation of bias-evaluating cancer treatment with epidemiological methods.* Acta Oncologica, 2006. **45**(4): p. 421-429.
111. Steineck, G., et al., *Symptom Documentation in Cancer Survivors as a Basis for Therapy Modifications.* Acta Oncologica, 2002. **41**(3): p. 244-252.
112. Wikman, A., *Reliability, Validity and True Values in Surveys.* Social indicators research, 2006. **78**(1): p. 85-110.
113. Wikman, A., *Att utveckla sociala indikatorer : en surveyansats belyst med exemplet arbetsmiljö*, u. Sverige. Statistiska centralbyrån, Editor. 1991, Diss. Stockholm : Univ.
114. Association, W.M., *World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects.* JAMA, 2013. **310**(20): p. 2191-2194.
115. Lang, T.A. and D.G. Altman, *Basic statistical reporting for articles published in Biomedical Journals: The "Statistical Analyses and Methods in the Published Literature" or the SAMPL Guidelines.* International Journal of Nursing Studies, 2015. **52**(1): p. 5-9.
116. von Elm, E., et al., *The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies.* Epidemiology, 2007. **18**(6): p. 800-804.
117. Lee, H., et al., *A Guideline for Reporting Mediation Analyses of Randomized Trials and Observational Studies: The AGReMA Statement.* Jama, 2021. **326**(11): p. 1045-1056.
118. Dahl, A.A., et al., *Long-term cervical cancer survivors on disability pension: a subgroup in need of attention from health care providers.* J Cancer Surviv, 2020.
119. Carlsen, K., et al., *Cancer and the risk for taking early retirement pension: a Danish cohort study.* Scandinavian Journal of Public Health, 2008. **36**(2): p. 117-25.
120. Nout, R.A., et al., *Long-Term Outcome and Quality of Life of Patients With Endometrial Carcinoma Treated With or Without Pelvic Radiotherapy in the Post Operative Radiation Therapy in Endometrial Carcinoma 1 (PORTEC-1) Trial.* Journal of Clinical Oncology, 2011. **29**(13): p. 1692-1700.
121. Nout, R.A., et al., *Quality of Life After Pelvic Radiotherapy or Vaginal Brachytherapy for Endometrial Cancer: First Results of the Randomized PORTEC-2 Trial.* Journal of Clinical Oncology, 2009. **27**(21): p. 3547-3556.
122. Nachreiner, N.M., R. Shanley, and R.G. Ghebre, *Cancer and treatment effects on job task performance for gynecological cancer survivors.* Work, 2013. **46**(4): p. 433-8.

123. Gami, B., et al., *How patients manage gastrointestinal symptoms after pelvic radiotherapy.* Alimentary Pharmacology & Therapeutics, 2003. **18**(10): p. 987-994.
124. Dahl, A.A., et al., *Employment Status and Work Ability in Long-Term Young Adult Cancer Survivors.* J Adolesc Young Adult Oncol, 2019. **8**(3): p. 304-311.
125. Haukenes, I., E.H. Farbu, T. Riise, and G.S. Tell, *Physical health-related quality of life predicts disability pension due to musculoskeletal disorders: seven years follow-up of the Hordaland Health Study Cohort.* BMC Public Health, 2014. **14**(1): p. 167.
126. Gómez-Molinero, R. and R. Guil, *Boosting return to work after breast cancer: The mediator role of perceived emotional intelligence.* Psycho-Oncology, 2020. **29**(11): p. 1936-1942.
127. Mateo-Rodríguez, I., et al., *Mediational Occupational Risk Factors Pertaining to Work Ability According to Age, Gender and Professional Job Type.* International Journal of Environmental Research and Public Health, 2021. **18**(3): p. 877.
128. Goedhard, R.G. and W.J.A. Goedhard, *Work ability and perceived work stress.* International Congress Series, 2005. **1280**: p. 79-83.
129. Peeters, M.C.W., J. de Jonge, P.P.M. Janssen, and S. van der Linden, *Work-home interference, job stressors, and employee health in a longitudinal perspective.* International Journal of Stress Management, 2004. **11**(4): p. 305-322.
130. Ennis, M., et al., *A comparison of statistical learning methods on the GUSTO database.* Statistics in Medicine, 1998. **17**(21): p. 2501-2508.
131. Halicka, M., M. Wilby, R. Duarte, and C. Brown, *Predicting patient-reported outcomes following lumbar spine surgery: development and external validation of multivariable prediction models.* BMC Musculoskeletal Disorders, 2023. **24**(1).
132. Diamond, G.A., *What price perfection? Calibration and discrimination of clinical prediction models.* J Clin Epidemiol, 1992. **45**(1): p. 85-9.
133. Rufibach, K., *Use of Brier score to assess binary predictions.* Journal of Clinical Epidemiology, 2010. **63**(8): p. 938-939.

.

APPENDIX 1 – STUDY I AND II

1800 gynaecological cancer survivors identified. Treated with external pelvic radiotherapy (EPRT) during 1991–2003

977 did not meet the first eligibility criteria
Deceased between treatment and this stage 497
Born before 1927 436
Could not understand/read Swedish 23
Had a recurrence 19
Had not received pelvic radiotherapy 2

An introduction letter was sent to the remaining 823 eligible survivors

92 did not want to participate
Reasons for non-participation:
No reason provided 29
Physical reason 21
Not reachable 17
Psychosocial reason 14
Psychological reason 9
Family members said no 2

731 survivors gave informed oral consent and were sent a questionnaire

81 non-responders
Sent back an empty questionnaire 29
Did not return the questionnaire 52

650 (79%) survivors returned a completed questionnaire

403 survivors were excluded
Reasons for exclusion
≥65 years old 363
Had undergone an ostomy 20
Missing social security number 7
≥30% missing values 7
Deceased before follow-up in 2008 6

247 survivors were included inStudy I and II

Figure 10.1: Flowchart of recruitment and selection of gynecological cancer survivors

Table 10.2: Self-reported gastrointestinal symptoms included in urgency syndrome and their estimated factor loading (syndrome intensity)

Factor loading	English translation	Original Swedish text
0.85	Sudden defecation urgency requiring lavatory	Har Du haft plötsligt påkomna avföringsträngningar toalettbesök
0.74	Immediate need to defecate	Har Du haft omedelbart behov av en toalett om Du behövt
0.71	Loose stools	Har Du haft lös avföring
0.69	Inability to hold stools for >5 minutes during urgency	Hur länge har Du kunnat hålla avföringen vid trängningar
0.61	Need to repeat defecation within one hour	Har Du återvänt till toaletten inom en timme efter avföring for att tömma"
0.40	Leakage of loose stools while awake	Har Du haft läckage av lös avföring när Du varit vaken"
0.35	Fecal leakage without warning despite pre-emptied bowel	Har Du utan förvarning läckt avföring i kläderna
0.34	Abdominal pain	Har Du haft smärtor i buken
0.32	Abdominal bloating	Har Du haft känsla av uppblåsthet i magen
0.27	Involuntary foul smelling flatulence	Har Du haft illaluktande gasavgångar som Du inte kunnat stoppa
0.26	Defecation into clothing without forewarning	Har Du utan förvarning tomt all avföring i kläderna
0.25	Involuntary unspecified flatulence	Har det hänt att Du inte har kunnat hålla kvar gas"
0.23	Anal itching	Har Du haft klåda vid ändtarmsöppningen, det senaste halvåret?
0.22	Leakage of loose stools while asleep	Har Du haft läckage av lös avföring, när Du sovit"
0.22	Anal pain	Har Du haft smarta i ändtarmsöppningen

In an earlier publication from our research team, self-reported radiation-induced symptoms were analyzed using modified exploratory factor analysis approach and factor loadings for each symptom, factor-specific factor-loading cutoffs and factor scores were estimated (Steineck et al. https://doi.org/10.1371/journal.pone.0171461)

Table 10.3: Age-adjusted Relative risks (RRs) and risk differences (RDs) (95% confidence intervals (CIs)) for disability pension (from the national register on disability pension) among survivors with one or more syndromes.

	Disability pension at 2-year follow-up			
	Relative risks[a] (95 % CI)		Risk differences[b] (95% CI)	
Syndrome[c]	**n = 247[d]**	**n = 243[e]**	**n = 247[d]**	**n = 243[e]**
Sum of syndromes[f]				
Three syndrome vs None	**2.7 (1.6 to 4.7)**	**2.8 (1.6 to 4.8)**	**30% (4% to 55%)**	**30% (5% to 56%)**
Two syndrome vs None	**2.5 (1.6 to 4.0)**	**2.6 (1.6 to 4.1)**	**23% (7% to 39%)**	**24% (8% to 40%)**
One syndrome vs None	1.3 (0.7 to 2.4)	1.3 (0.7 to 2.4)	5% (-9% to 18%)	5% (-9% to 19%)
Urgency vs No urgency syndrome	**2.0 (1.3 to 3.0)**	**2.0 (1.4 to 3.0)**	**17% (4% to 29%)**	**17% (5% to 30%)**
Leakage vs No leakage syndrome	**2.0 (1.4 to 3.0)**	**2.1 (1.4 to 3.0)**	**18% (5% to 31%)**	**18% (5% to 31%)**
Blood vs No blood discharge syndrome	**2.1 (1.4 to 3.1)**	**2.1 (1.4 to 3.1)**	**23% (4% to 41%)**	**23% (4% to 42%)**
Excessive gas discharge vs No excessive gas discharge	**1.7 (1.2 to 2.6)**	**1.8 (1.2 to 2.7)**	11 % (-4 to 26%)	11 % (-4 to 26%)
Excessive mucus discharge vs No excessive mucus discharge	1.1 (0.7 to 1.7)	1.1 (0.7 to 1.7)	1 % (-12% to 14%)	1 % (-12% to 14%)

[a b] Age-adjusted RR and RD (95% CI) obtained from log-binomial regression analyses using *Syndrome* and *Age (in years)* as independent variables, '*None/No*' level of *Syndrome* was used as a reference. [c] Self-reported symptoms were used to build *Syndromes*. [d] Survivors alive at follow-up in 2008. [e] Excluding survivors who died within the 2-years of follow-up (between 2008 and 2010) [f] Survivors classified as having several syndromes or one or none. Bold **numbers** indicate a statistically significant association at 5% level of significance.

Table 10.4: Single-item questions on different aspects of self-reported quality of life (QoL) and answer alternatives from the postal questionnaire

English translation		Original Swedish text	
Global quality of life			
How has your quality of life been in the last 6 months?		Hur har Din livskvalitet varit, det senaste halvåret?	
1 = No quality of life at all	7 = Best possible quality of life	1 = Ingen livskvalitet alls	7 = Bästa möjliga livskvalitet
Global physical health			
How has your physical health been in the last 6 months?		Hur har Din kroppsliga hälsa varit det senaste halvåret?	
1 = Worst imaginable health	7 = Best imaginable health	1 = Sämsta tänkbara hälsa	7 = Bästa tänkbara hälsa
Physical strength (condition)			
How has your physical strength (condition) been in the last 6 months?		Hur har Din kroppsliga ork (kondition) varit, det senaste halvåret?	
1 = No strength	7 = Best imaginable strength	1 = Ingen ork	7 = Bästa tänkbara ork
Depressed or feeling sad			
Have you felt down or depressed in the last 6 months?		Har Du känt Dig nedstämd eller deprimerad, det senaste halvåret?	
1 = Never	7 = All the time	1 = Aldrig	7 = Hela tiden
Worry or anxiety			
Have you felt worry or anxiety in the last 6 months?		Har Du känt oro eller ångest, det senaste halvåret?	
1 = Never	7 = All the time	1 = Aldrig	7 = Hela tiden
Psychological wellbeing			
How has your mental wellbeing been in the last 6 months?		Hur har Ditt psykiska välbefinnande varit, det senaste halvåret?	
1 = No wellbeing	7 = Best possible wellbeing	1 = Inget välbefinnande	7 = Bästa tänkbara välbefinnande
Having meaning in life			
Has your life felt meaningful in the last 6 months?		Har Ditt liv känts meningsfullt, det senaste halvåret?	
1 = Never	7 = All the time	1 = Aldrig	7 = Hela tiden
Self-esteem			
How has your self-esteem been in the last 6 months?		Hur har Din självkänsla varit, det senaste halvåret?	
1 = No self-esteem	7 = Best imaginable self-esteem	1 = Ingen självkänsla	7 = Bästa tänkbara självkänsla
Satisfied with sleep			
How satisfied have you been with your sleep in the last 6 months?		Hur nöjd har Du varit med Din situation, det senaste halvåret? Sömn	
1 = Not at all satisfied	7 = Completely satisfied	1 = Inte alls nöjd	7 = Helt nöjd
Satisfied with concentration			
How satisfied have you been with your concentration in the last 6 months?		Hur nöjd har Du varit med Din situation, det senaste halvåret? Koncentration	
1 = Not at all satisfied	7 = Completely satisfied	1 = Inte alls nöjd	7 = Helt nöjd
Satisfied with memory			
How satisfied have you been with your memory in the last 6 months?		Hur nöjd har Du varit med Din situation, det senaste halvåret? Minne	

APPENDIX 2 – STUDY III

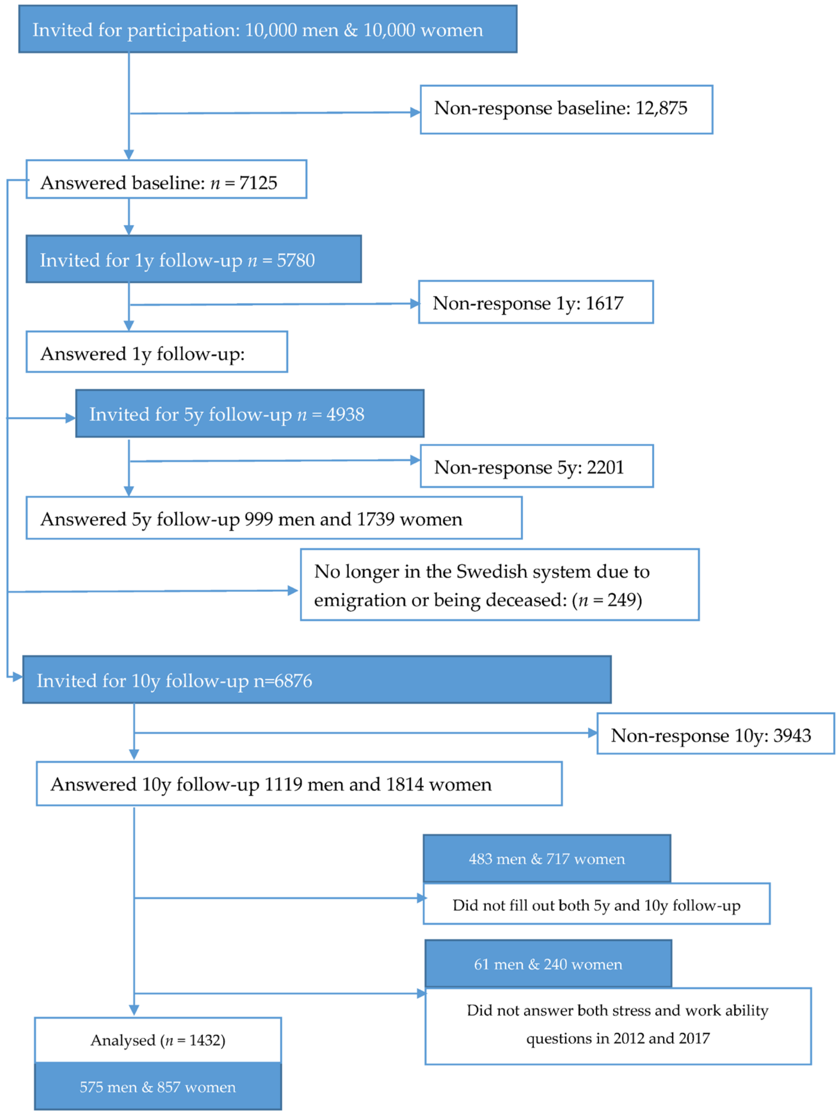

Figure 11.1: Flowchart of inclusion of the Work Ability in Young Adults (WAYA) cohort. Ref: van Schaaijk, A. et al (2020) Mediating Factors for the Relationship between Stress and Work Ability. https://doi.org/10.3390/ijerph17072530

APPENDIX 3 –STUDY IV

Table 12.1: English translation of the work environment questions included in Model 1 with 7 predictors.

Nr	Variable name	Question in the survey (explanation)
1	WAS	Work ability score. Assume that your work ability at its best has a value of 10 points. How many points would you give your current work ability?” with a score range of 1–10.
2	BArbePens	Do you think you will be able to work until the normal retirement age in your profession?
3	BetalaUt	In the last 12 months have you received training on company time?
4	BFHVBedom	Have occupational health services assessed your work situation in any other ways? For example, in connection with your visit to occupational health services or by telephone?
5	SovaGott	Do you think you get enough sleep?
6	VilaGott	Besides sleep, do you think you get adequate time for resting and relaxation between working days?
7	SANDOVE	Due to health reasons, have you considered in the last year changing work responsibilities, changing present work responsibilities, changing employer or becoming self-employed, or reducing working hours?

Table 12.2: English translation of the work environment questions, demographical, and other variables included in Model 2 (13 predictors)

Nr.	Variable name	Question in the survey or explanation
1	WAS	Work ability score. Assume that your work ability at its best has a value of 10 points. How many points would you give your current work ability?” with a score range of 1–10.
2	BArbePens	Do you think you will be able to work until the normal retirement age in your profession?
3	BetalaUt	In the last 12 months have you received training on company time?
4	BFHVBedom	Have occupational health services assessed your work situation in any other ways? For example, in connection with your visit to occupational health services or by telephone?
5	SovaGott	Do you think you get enough sleep?

6	VilaGott	Besides sleep, do you think you get adequate time for resting and relaxation between working days?
7	SANDOVE	Due to health reasons, have you considered in the last year changing work responsibilities, changing present work responsibilities, changing employer or becoming self-employed, or reducing working hours?
8	Sex	Sex of the respondent, female or male
9	Civil	Marital status Single, Married or registered partner, Divorced or widowed
10	FamTypF	Family composition. Single parent, Children <18 years of age, No children or children >18 years of age,
11	Sun2000niva	The educational level was revised in 2000. Primary, Lower secondary, and Upper Secondary Post-secondary non-tertiary, Short-cycle tertiary, Bachelor or equivalent, Master or equivalent, Doctoral or equivalent
12	SSYK2	Swedish Standard Classification of Occupations 1996 (two-digit level with 27 groups). SSYK is based on the International Standard Classification of Occupations, ISCO-88. E.g., Corporate managers, Life science and health professionals, etc.
13	SektorKodSreg	The sectoral affiliation of the company/employer. Central government, State-owned enterprises, Primary municipal administration, County councils, Other public institutions, Limited liability companies (not publicly owned), Other enterprises (not publicly owned), State-owned enterprises and organizations, municipally owned enterprises and organizations, and other organizations.

Table 12.3: English translation of the work environment questions, demographical, and other variables included in Model 3.

Nr.	Variable name	Question in the survey (explanation)
1	WAS	Work ability score. Assume that your work ability at its best has a value of 10 points. How many points would you give your current work ability?" with a score range of 1–10.
2	SSYK2	Swedish Standard Classification of Occupations 1996 (two-digit level with 27 groups). SSYK is based on the International Standard Classification of Occupations, ISCO-88. E.g., Corporate managers, Life science and health professionals, etc.
3	ProblemP	In your job do you ever come into contact with people who are seriously ill or people with serious problems?
4	KopplaAv	Does it happen that you cannot dismiss your job from your thoughts when you are off work?
5	Sun2000niva	The educational level was revised in 2000. Primary, Lower secondary, and Upper Secondary Post-secondary non-tertiary, Short-cycle tertiary, Bachelor or equivalent, Master or equivalent, Doctoral or equivalent
6	BArbePens	Do you think you will be able to work until the normal retirement age in your profession?
7	PrataPau	In the main can you take short breaks at any time to talk?
8	KroppsAn	Is your work physically demanding?
9	Overtide	Do you sometimes have so much work to do that you have to skip lunch, work late, or take work home with you?
10	Sovsvari	Have you during the last three months had a hard time sleeping because thoughts about your work keep you awake?
11	VilaGott	Besides sleep, do you think you get adequate time for resting and relaxation between working days?
12	Missnojd	Generally speaking, I am very dissatisfied with my work
13	TungLatt	How do you experience your work? Strenuous heavy work
14	StorBel	Does your work involve a heavy workload, i.e., too much to do?
15	Pafresta	How do you experience your work? Mentally stressful work
16	TrottHog	Have you during the last three months been tired and listless?
17	JobbStim	Is your work interesting and stimulating?
18	IntePrat	Is your work sometimes so stressful that you do not have time to talk or even think of anything other than work?
19	UppRor	Does your work involve unilateral repetitive movements?
20	Lyft1525	Do you have to lift at least 15 kg several times a day?

21	VriderOf	Does it happen at work that you bend or turn in the same way many times per hour for several hours in one day?
22	SovaGott	Do you think you get enough sleep?
23	KroppStor	Is your work physically demanding and involves a heavy workload, i.e., too much to do?
24	MyckLite	How do you experience your work? Far too much to do
25	Meningsf	How do you experience your work? Extremely meaningless work
26	BundFrit	How do you experience your work? Constrained and unfree
27	TaktBest	Is it possible for you to set your work tempo?
28	SANDOVE	Due to health reasons, have you considered in the last year changing work responsibilities, changing present work responsibilities, changing employer or becoming self-employed, or reducing working hours?
29	Enformig	How do you experience your work? Monotonous work
30	Sex	Sex of the respondent, female or male
31	ChefUpps	Does your superior (boss) ever express appreciation for your work?
32	BestTide	Is it possible for you to decide on your own when various tasks are to be done (for example, by choosing to work a bit faster on some days and taking it easier on other days)?

These thirty-two variables were marked by Boruta as relevant, an ML wrapper method built around the random forest classifier algorithm to perform variable selection. Initially, forty-eight candidate variables were input in Boruta, encompassing various work environments and demographical variables.

Table 12.4: English translation of the sixteen variables from the Work Environment Survey, marked by Boruta as irrelevant.

Nr.	Variable name	Question in the survey (explanation)
1	Stallnin	How do you experience your work? Strenuous work postures
2	FamTypF	Family composition
3	HotOmVal	Are you exposed to violence or the threat of violence in your work?
4	CheferSt	Can you receive support and encouragement from your superiors when your work becomes troublesome?
5	KamratSt	Can you receive support and encouragement from your fellow workers when your work becomes troublesome?
6	BestUppl	Are you involved in planning your work (for example, what is to be done, how it is to be done, or who is to work with you)?

7	SektorKodSreg	The employer/company's sector code
8	BetalaUt	In the last 12 months have you received training (education and learning) on company time?
9	Koncentr	Does your work require your undivided attention and concentration?
10	AndraUpp	Do other persons express appreciation for your work (e.g., fellow workers, patients, customers, clients, passengers, students)?
11	SvarEnkl	How do you experience your work? Tasks too difficult
12	AtgBrist	Are deficiencies in the work environment taken care of within a reasonable time (as soon as you feel you can expect)?
13	Systemat	Questions concerning systematic work environment management. Are such systematic activities currently ongoing at your workplace?
14	BFHVBedom	Have occupational health services assessed your work situation in any other ways? For example, in connection with your visit to occupational health services or by telephone?
15	Civil	Marital status
16	Inflytan	How do you experience your work? Too little influence

These sixteen variables were marked by Boruta as irrelevant, an ML wrapper method built around the random forest classifier algorithm to perform variable selection. Initially, forty-eight candidate variables were input in Boruta, encompassing various work environments and demographical variables.