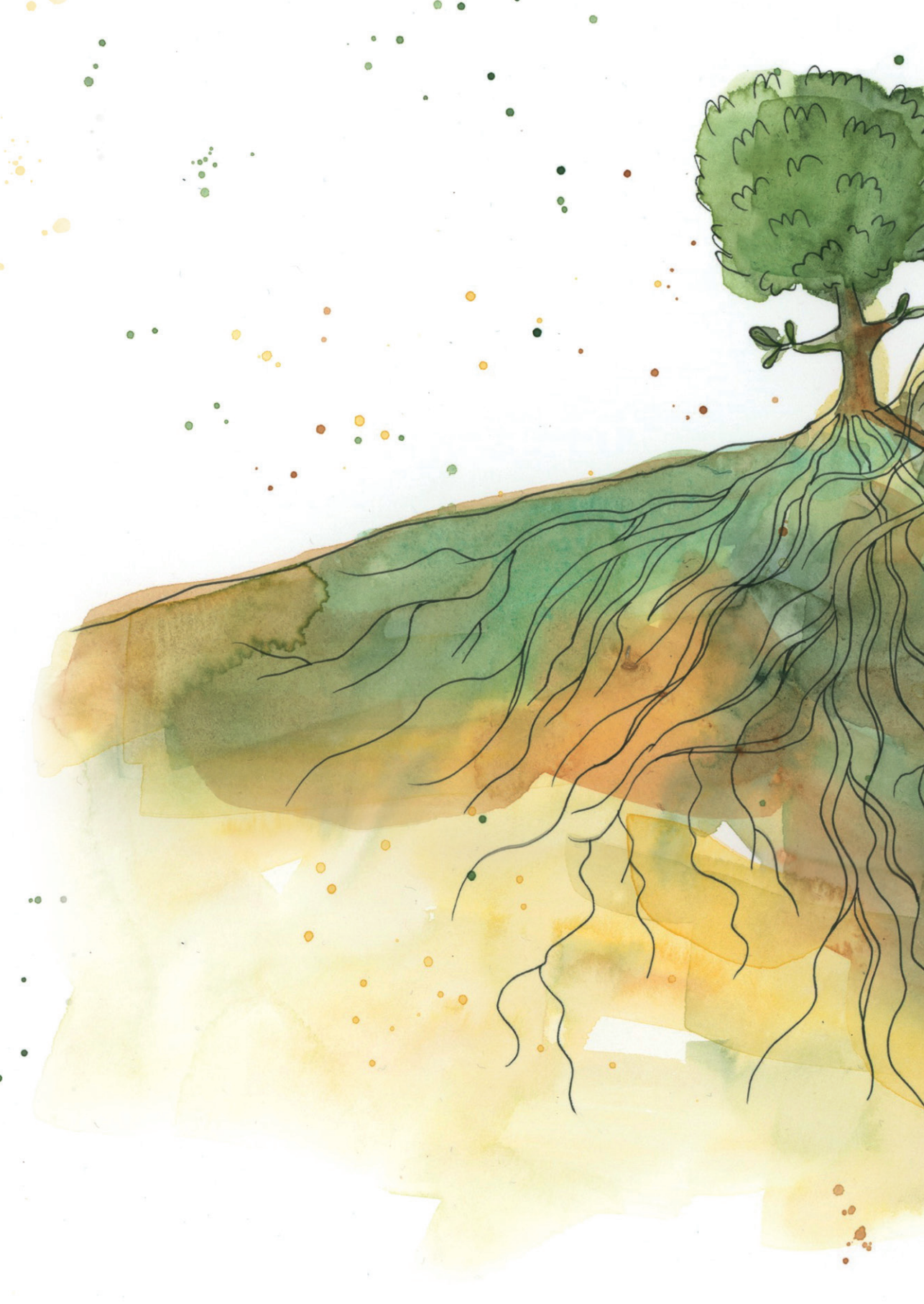
TWISTS AND TURNS

Sex and Gender Differences in the Illness Trajectories of Common Somatic Symptoms

ARANKA VIVIENNE BALLERING



TWISTS AND TURNS

Sex and Gender Differences in the Illness Trajectories of Common Somatic Symptoms

ARANKA VIVIENNE BALLERING

Colophon
This thesis was supported by the University Medical Center Groningen, the Graduate School of Medical Sciences, Research Institute SHARE and the University of Groningen.

Cover design and layout: © evelienjagtman.com
Print: Ridderprint | www.ridderprint.nl

This research was supported by the ZonMw Gender and Health Program (849200013 and 849700005), the ZonMw COVID-19 Program (10430302110002), and NWO Talent Program Vici (Vi.C.191.021).

rijksuniversiteit
groningen

TWISTS AND TURNS

Sex and Gender Differences in the Illness Trajectories of Common Somatic Symptoms

Proefschrift

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. ir. J.M.A. Scherpen
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

woensdag 3 april 2024 om 12.45 uur

door

Aranka Vivienne Ballering

geboren op 23 september 1994
te Alphen aan den Rijn

Promotor
Prof. dr. J.G.M. Rosmalen

Copromotor
Dr. T.C. Olde Hartman

Beoordelingscommissie
Prof. dr. R.A. Schoevers
Prof. dr. N.D. Scherpbier
Prof. dr. H.M. Den Ruijter

Thesis Outline

Chapter 1	General introduction	9
Etiology of common somatic symptoms		
Chapter 2	The association of sex, age and FKBP5 genotype with common somatic symptoms: A replication study in the Lifelines cohort study. *Journal of Psychosomatic Research 2021;147:110510*	27
Chapter 3	Discordance between adolescents and parents in functional somatic symptom reports: Sex differences and future symptom prevalence. *Journal of Youth and Adolescence 2023;52:2182–2195*	39
Prevalence and persistence of common somatic symptoms		
Chapter 4	Gender and sex independently associate with common somatic symptoms and lifetime prevalence of chronic disease. *Social Science & Medicine 2020;253:112968*	67
Chapter 5	Female sex and femininity independently associate with common somatic symptom trajectories. *Psychological Medicine 2022;52(11):2144-2154*	95
Chapter 6	Persistence of somatic symptoms after COVID-19 in the Netherlands: an observational cohort study. *The Lancet 2022;400(10350):452-461*	131
Primary care help-seeking for common somatic symptoms		
Chapter 7	Sex and gender differences in primary care help-seeking for common somatic symptoms: a longitudinal study. *Scandinavian Journal of Primary Health Care 2023;41(2):132-139*	161
Diagnostics in primary care for common somatic symptoms		
Chapter 8	Sex differences in incidence of respiratory symptoms and management by general practitioners. *Family Practice 2020;37(5):631-636*	183
Chapter 9	Sex differences in the trajectories to diagnosis in patients presenting with common somatic symptoms in primary care: an observational cohort study. *Journal of Psychosomatic Research 2021;149:110589*	199

Chapter 10	Differences between women and men are present in the rate of diagnosed diseases after a diagnostic intervention is conducted in primary care. *Journal of the American Board of Family Medicine 2022;35(1):73-84*	225
Chapter 11	Sex and gender-related differences in COVID-19 diagnoses and SARS-CoV-2 testing practices during the first wave of the pandemic: the Dutch Lifelines COVID-19 cohort sudy. *Journal of Women's Health 2021;30(12):1686-1692*	245
Chapter 11A	Response to Rossato et al. *Journal of Women's Health 2022;31(6):896-898*	267
Adequate conceptual assessment and implementation of knowledge		
Chapter 12	How to ensure inclusivity in large-scale population cohort studies? Lessons learned with regards to including and assessing sex, gender and sexual orientation. *Archives of Sexual Behavior 2023;52,2163–2172*	275
Chapter 13	Building a bridge between multidisciplinary insights and practice: The development of an e-learning for internal residents about sex, gender and persistent somatic symptoms. *Journal of Psychosomatic Research 2022;155:110739*	295
Chapter 14	General discussion	303
Appendices		
Summary		337
Samenvatting		345
About the author		353
List of publications		357
Dankwoord		363
Research Institute SHARE		371

CHAPTER 1

General introduction

We all experience symptoms, such as headaches, muscle aches, general tiredness and lower back pain every once in a while. In any given week, approximately 80% of the general population experiences these type of symptoms, indicating these are rather common.[1] These symptoms may be attributed to a clear organic cause, but frequently no underlying disease can be established. In the vast majority of people these symptoms spontaneously appear and subsequently disappear as well.[2] Generally, common somatic symptoms are self-limiting. In some people the symptoms may persist and may be accompanied by limitations in daily functioning, risk of iatrogenic harm due to unnecessary interventions, and high levels of personal suffering.[3]

The illness trajectory for common somatic symptoms, starting with the first occurrence of symptoms is a complex, uncertain and iterative process. In this dynamic process multiple decisions are made, resulting from many interactions between patients, healthcare professionals and other relevant actors.[4,5] Throughout the illness trajectory for common somatic symptoms sex and gender differences can be observed. Women were found to experience more frequent, more persistent and more intense somatic symptoms than men.[6-8] Similarly, women were previously found to be less satisfied with the primary care they receive for their somatic symptoms than men.[9] This thesis studies multiple aspects of the illness trajectories related to common somatic symptoms in depth, and explores potential sex and/or gender differences and implications hereof within these trajectories.

First, the history of sex and gender sensitive medicine (SGSM) is briefly introduced. Second, the concepts of sex and gender and especially their application in epidemiology are described in depth. Last, illness trajectories and the outline of this thesis are discussed in more detail.

A brief timeline of sex and gender sensitive medicine

Hippocrates is considered the father of modern medicine. Hippocratic ideas on medicine, especially those on women's health, formed the basis for long-maintained misconceptions. The most prominent idea within Hippocratic gynecology was that of women having a 'wandering womb', with each direction of movement related to a unique set of symptoms (e.g., headache occurred when the uterus moved towards the head). This idea was maintained by some influential physicians up until the late Medieval period and was accompanied by the belief that the uterus was moving, because it

was in search for a child to carry.[10,11] A favored treatment of this 'mobile uterus' were amulets and charms worn by women.[12] Some of these age-old stereotypical ideas regarding women's health have persisted over time. The uterus largely remained the scientific focus of women's health even up until late in the nineteenth century. By then women's health was generally reduced to the female disease of hysteria, which was a catch-all diagnosis for symptoms experienced by women.[13] Hysteria was provided as an explanation for shortness of breath, fainting, a decreased but also an increased female libido, and many more somatic experiences of women. The appropriate treatment for hysteria was, again, centered around the uterus and included drastic measures, such as hysterectomies. Hysteria was only removed from the Diagnostic and Statistical Manual of mental disorders (DSM) in 1980.

Fortunately, medical science and medicine broadened their scope. In the past century women's health was still considered women's business. Starting from the 1970s, medicine became gradually more sex and gender-sensitive, as the medical community realized that both biological differences between women and men (i.e., sex differences), but also psychosocial differences (i.e., gender differences) were relevant to health and illness. First, an important milestone was the introduction of the Bem Sex Role Inventory (BSRI) as an assessment of gender in 1974.[14] The BSRI quantified then-relevant feminine and masculine traits as defining distinct dimensions rather than opposing femininity to masculinity. However, the BSRI is criticized for its operationalization of gender by using stereotypical personality traits to define femininity, for example by traits such as being 'shy', 'gullible' and 'soft-spoken', or 'self-reliant', and masculinity by 'dominant' and 'ambitious', respectively, as well as for the lack of validation.[15,16] Despite the criticism, the BSRI allowed for a bidimensional approximation of gender to be integrated in health research, showcasing researchers' efforts to include gender next to sex in health research. Although sociology had previously established the significance of gender for health and health research, the objectivist ideas of biomedicine remained hardly compatible with the constructionist ideas of sociology.[17] Biomedical research adhered to the notion of a singular, objective reality that is universally applicable to all individuals, while sociology recognized the construction of reality by many individuals via subjective experiences. The quantification of gender via the BSRI bridged these differences and eased the inclusion of gender as a socially-constructed concept in medicine.

Second, the movement that aimed to establish and mainstream SGSM remained narrowly focused on female-specific health conditions, related to reproductive medicine and obstetrics.[18] It was rather sex-centered (i.e., it was centered around the female biology and diseases of the female body), with little room for gender sensitive medicine (i.e.,

sensitivity to the sociocultural implications of being a woman on health). The movement swiftly grew beyond that, ultimately seeking improved recognition and acknowledgement for the different symptoms that men and women experience for similar underlying diseases. It grew into a movement that was wary of male-b(i)ased medicine, fueled by, among other things, the ideas of feminist philosophers such as Simone de Beauvoir.[18]

Over time, multiple methods beyond the BSRI were developed to operationalize and quantify gender.[19] This further aided in overcoming difficulties related to including a socially-constructed concept in the objectivist dogma of medicine, although some friction remained. Subsequently sex and gender-sensitivity gained ground in (bio) medical science and epidemiology, raising awareness for how both sex and gender affect people's health. Currently, even the objectivist dogma of medicine is subject to change. This allows for more constructionist views to enter the (bio)medical realm, resulting in a more effective and comprehensive incorporation of gender in health research and thus personalized medicine. Nowadays, SGSM aims to improve healthcare for all variations in sex and gender, not merely female patients and women,[20] and finds itself as being an integral part of personalized medicine.[17,18] To this end, consideration of sex and gender in health research is now frequently required by funders and academic journals. This is not surprising as sex and gender sensitivity in research allows for rigor in and reproducibility of the scientific process, consequently resulting in more valid and generalizable research outcomes.[18,21]

Sex and gender – the who, the why, and the what

Sex and gender in healthcare are currently hotly debated topics.[22] Although one could argue this is advantageous for the further development of SGSM, the discussions on this topic frequently lack nuance. For example, the concepts of sex and gender are often conflated or equated especially in biomedical and epidemiological research settings,[23] all women supposedly have similar illness trajectories with them consistently disadvantaged compared to men, and gender is usually only considered relevant for people identifying as transgender or gender-diverse. Misconceptions, including but certainly not limited to these three aforementioned examples, result in a polarized debate regarding the role of sex and gender in healthcare.

Sex and gender are in a continuous dialogue shaping each other, while being subjected to societal norms. Nevertheless, the two concepts are distinctly different and we do not necessarily consider one as the consequence of the other. So, before further explicating the role of sex and gender in healthcare and in illness trajectories, the

conceptual difference between sex and gender should be clarified. On the one hand, sex encompasses the biology of male and female bodies. It refers to the biological features and aspects, such as physiology, anatomy, gene expression, and hormone levels and function, that define female and male bodies.[24] Sex is usually assigned at birth. Sex is not a dichotomy, although often considered as such, but rather a continuum ranging from male to female and vice versa. This implies the presence of intersex variations within bodies as well: bodies with intersex variations do not conform to the archetypical medical and social ideas of what constitutes a male or female body.[8] However, where the boundaries of the male and female body lie, is not yet clearly defined. Whether we should define these boundaries at all and whether this definition is desired or even necessary, or merely a deeply rooted, ingrained human need for categorization, is currently topic of debate.[25]

Gender, on the other hand, refers to a multidimensional and socioculturally-constructed concept, which is strongly dependent on place and time. Although gender is frequently regarded as the psychosocial equivalent of sex, no causality should be conferred from that notion as it may unintentionally reinforce the idea that gender is a sole consequence of sex. Gender encompasses among others, the dynamic embodiment of identities, behaviors and roles within a given society.[26] The exact dimensions of gender are subject to continued debate.

An increasing body of evidence shows that illness trajectories for common somatic symptoms are affected by (the degree of adherence to) socially prescribed norms and experiences of 'being a man' and/or 'being a woman',[27-29] regardless of the exact interpretation of gender dimensions. This thesis distinguishes between four dimensions of gender that we consider most relevant for epidemiological and healthcare-related research.

I) *Gender identity* refers to how individuals see themselves, based on internal(ized) and personal feelings, on the continuum of man to woman or beyond that. One's gender identity is fluid, dependent on societal and cultural norms, and over time.[24,26] People whose gender identity aligns with their biological sex, which is usually assigned at birth, are considered cisgender, whereas people whose gender identity does not fully align with their biological sex are considered transgender or gender-diverse people.[25]

II) *Gender roles* are the roles individuals (are expected to) take upon themselves within a society. These are reflective of behavioral norms and mores imposed upon people based on them being considered a man, woman or transgender or

genderdiverse.[25,30] This results in different expectations and opportunities for people. These expectations and opportunities are found in different domains, such as the workplace, domestic settings and the social space.[22,26] For this thesis, we consider gender expression to be a part of gender roles, although it is argued that gender expression may be a distinct dimension of gender.[25] Expression of gender involves individuals' demonstration and enactment via, among others, behaviors, activities, dress codes, mannerisms, appearance and societal opinions.[25] Although with gender expression the focus is on the individual, gender expression is strongly subjected to the behavioral norms and mores imposed upon people by the collective. Recently, innovative methods of quantifying people's gender roles to incorporate these in epidemiological studies have been developed,[31,32] with some specifically focusing on effectively incorporating the time, place and society-bound nature of gender roles.[8]

III) *Gender relations* refer to how individuals interact with, are understood, and are treated by others based on their gender.[25] Relations are inherently reciprocal and their nature is not solely defined by the individual, but is created upon interaction in various social settings, such as the work setting, family settings and social gatherings.

IV) *Institutionalized gender* refers to a hierarchy in terms of power between genders in among others, political, religious, medical and social institutions, in a given time and society.[26] These institutions shape and frequently reproduce gendered norms, concomitantly justifying these on a societal level.[25]

Notably, gender has an inherent interactive aspect to it with its embodiment being subject to ever-changing societal norms. Gender consciously or unconsciously plays a role in all social interactions, relationships and institutions, allowing for reproduction and potential reinforcement of existing norms and mores regarding hierarchies between genders.[25,33] The proneness of gender to hierarchical pressures exemplifies a mechanism via which gender may intersect with other social determinants, such as socioeconomic status and class, affecting health.[34]

Up until recently, gender has been largely neglected in health research,[23] with previous research mainly focusing on potential sex differences in health. However, both sex and gender uniquely associate with a plethora of health outcomes.[7,8] A classic example hereof is osteoporosis: differences in biological parameters, including bone density and hormone levels, are known to contribute to the female preponderance of osteoporosis.[35] Traditional feminine gender roles also discourage participating

in sports that involve heavy weightlifting and may promote caloric restriction, contributing to an increased risk of osteoporosis in women compared to men.[36] Deeper insights about the contributions of gender and its concomitant dimensions to health, next to sex, will aid more personalized medicine.

Illness trajectories of common somatic symptoms

An illness trajectory starts when a person notices a bodily sensation that is interpreted to be "wrong", or in other words as a symptom. An illness trajectory only ends if the symptom is resolved and the care that is provided for the symptom is considered to be finished.[5] During an illness trajectory, which may last from a mere few hours to many years, multiple so-called "critical junctions" occur.[37] These are turning points involving an experience or event, or interdependent sequence of events, that have a potentially far-fetching impact on the patient's health and healthcare experience.[38] Critical junctions include, but are not limited to interpreting the severity of a symptom, persistence of a symptom, seeking care for symptoms, being provided with diagnostic interventions for the symptom and being provided with a diagnosis or an explanation, or the absence hereof, for the symptom. The latter critical junction, being provided with a diagnosis, marks the moment an illness becomes a disease.[39] Semantically, this critical junction would transform an illness trajectory into a disease trajectory, but in practice an individual's illness (i.e., the experience of ill health) does not stop or automatically worsen or improve when a disease is diagnosed.[40] A diagnosis legitimizes symptoms, may offer a prognosis and treatment strategies, and may become part of people's social identities.[41] Yet, a diagnosis may be viewed as a double-edged sword with on the one hand positive aspects such as legitimization and validation of symptoms, but on the other hand far-fetching biopsychosocial consequences that have implications far beyond the physician's consultation room.[5] Nevertheless, the absence of a diagnosis that adequately explains the somatic symptoms has far-fetching consequences as well. These consequences may include, but are not limited to concerns about people's ideas about the legitimacy of the somatic symptoms or feeling a lack in sense of belonging to a disease-specific patient group.

As illustrated above, the complex and iterative nature of illness trajectories results in critical junctions having implications for the progression of the remainder of the illness trajectory. An illness trajectory, including its critical junctions, is embedded in a real-life sociocultural context, allowing for a multitude of biological and psychosocial factors, including sex and gender, to partly shape it. Differences in the frequencies

with which specific critical junctions occur exist between women and men,[6-8,42] also in the illness trajectory's phase beyond a diagnosis.[22] Additionally, sex and gender differences in how critical junctions are substantiated occur as well, as it was recently shown that the interactions between healthcare professionals and patients varied depending on their gender.[43]

Aim and outline of this thesis

Although previous epidemiological studies have indicated the occurrence of sex differences in (parts of) illness trajectories of a variety of health conditions, these studies frequently have methodological, design, and interpretative limitations. First, previous studies frequently oversimplify differences between male and female patients, disregarding gender-related factors and intersex conditions. This is most likely due to either a lack of awareness about the importance of gender, or due to the lack of adequate epidemiological measures for gender.[24] Second, many previous studies, such as those focusing on help-seeking for somatic symptoms, were designed to be conducted in patient populations. This is problematic as it automatically excludes people who do not seek help for their illness or those who do not consider themselves patients.[42] Third, results from multiple previous studies have interpreted differences between women and men in illness trajectories instantly as gender inequalities without providing nuance regarding these differences. Some differences between men and women in illness trajectories may be justified.

Despite the wealth of evidence that exists regarding the relevance of sex and gender differences in health in general, knowledge on these differences in the illness trajectory of common somatic symptoms is lacking. Therefore, the overall aim of this thesis is to gain in-depth insights into whether and how sex and gender are associated with the illness trajectories of common somatic symptoms, while taking into account the multifaceted nature of sex and gender.

This thesis is structured according to the illness trajectory of common somatic symptoms, with critical junctions demarcating the thesis' different sections (***Figure 1***). Although ***Figure 1*** visualizes the illness trajectory as a linear process, this is an oversimplification as in reality an illness trajectory is a complex, uncertain and iterative process, which not necessarily ends when a diagnosis or explanation is provided for the experienced symptoms.[5] It should be noted that more critical junctions may occur in an illness trajectory than depicted in ***Figure 1***.

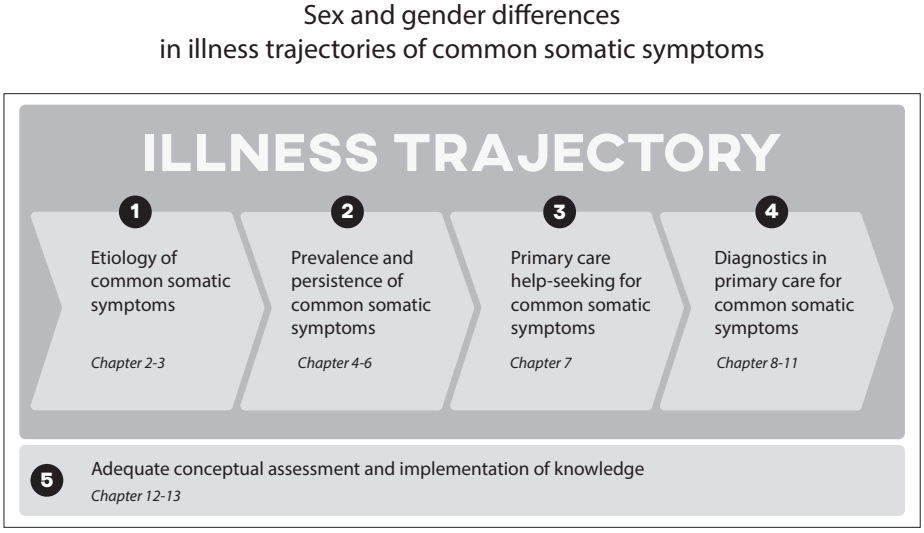


Figure 1: Structure of this thesis, modelled by the illness trajectory of people with common somatic symptoms.

1. Etiology of common somatic symptoms

First, this thesis aims to assess whether sex and/or gender differences affect the etiology of common somatic symptoms. The thesis starts by studying the biological etiology of common somatic symptoms using data from the large scale Dutch Lifelines Cohort Study. In *Chapter 2* we assessed whether sex differences occur in the genetic contribution to common somatic symptoms in adults. *Chapter 3* examines aspects of the social etiology of common somatic symptoms using data from the general population cohort Tracking Adolescents' Individual Lives Survey (TRAILS). We studied longitudinally whether sex and gender differences in common somatic symptom experience during young adulthood are predicted by the incongruency between parental reports and children's own report of somatic symptoms in adolescence.

2. Prevalence and persistence of common somatic symptoms

The second section of this thesis comprises empirical studies that aim to answer whether sex and gender independently associate with common somatic symptoms and symptom progression in women and men. To assess the unique cross-sectional associations of sex and gender with the prevalence of common somatic symptoms, we describe the development of a methodology to calculate a novel data-driven measure for gender within the Lifelines Cohort Study in *Chapter 4*. In addition, we cross-sectionally assess whether sex and gender uniquely associate with the prevalence of common somatic symptoms. Then, in *Chapter 5*, we longitudinally examine whether symptom severity

develops differently over time in women and men using Lifelines data. In *Chapter 6* the development of the severity of 23 symptoms surrounding a COVID-19 diagnosis was assessed for women and men separately by making use of longitudinal Lifelines COVID-19 Cohort Study.

3. Primary care help-seeking for common somatic symptoms
In the third section, we examine whether potential differences in the frequency of primary care help-seeking for common somatic symptoms associate with sex and gender. In order to study this, we linked a patient registry and a general population cohort, namely the NIVEL primary care database and the Lifelines Cohort Study in *Chapter 7*.

4. Diagnostics in primary care for common somatic symptoms
The studies in the fourth section of this thesis focus on the potential occurrence of sex differences in the diagnostic interventions, predominantly provided in primary care. In *Chapter 8* we assess the occurrence of sex differences in the incidence of respiratory symptoms and the management hereof. This study is conducted in a retrospective cohort study, using data from the world's oldest practice-based research network FaMe-Net, in which all morbidity presented to participating GPs is systematically recorded in episodes of care. In *Chapter 9* we expanded hereon, by assessing whether male and female patients are provided with similar diagnostic interventions and whether potential differences in the provision of diagnostic interventions affect patients' final diagnosis. Following along the lines of diagnostic interventions, we examine the effectiveness of diagnostic interventions in primary care in both women and men in *Chapter 10*. Thereafter, we focus on diagnostic interventions during the COVID-19 pandemic. We explore whether sex and gender-related factors associate with SARS-CoV-2 testing practices and COVID-19 diagnoses during the first wave of the pandemic in *Chapter 11* using Lifelines data. We supplement this with an analysis of sex differences in hospitalization due to COVID-19.

5. Adequate conceptual assessment and implementation of knowledge
In the fifth section of this thesis we identify the pitfalls we encountered when studying sex and gender differences in health when using data derived from large-scale population cohort studies in *Chapter 12*. We also propose concrete strategies to overcome these caveats. In *Chapter 13* we describe the development of an e-learning course about sex and gender differences in illness trajectories. Last, in *Chapter 14* we provide an overview of this thesis' main findings, while acknowledging the strengths and limitations of the studies included in this thesis. We also move beyond the research results and place these in the context of previous research and current societal developments regarding sex and gender. Finally, we will discuss implications for future research and clinical practice.

References

1. Acevedo-Mesa A, Tendeiro JN, Roest A, Rosmalen JGM, Monden R. Improving the measurement of functional somatic symptoms with item response theory. *Assessment*. 2020;28(8):1960-1970. doi: 10.1177/1073191120947153.
2. Chaabouni A, Houwen J, Walraven I, et al. Patients' characteristics and general practitioners' management of patients with symptom diagnoses: A retrospective cohort study. *Journal of the American Board of Family Medicine*. 2023(36):477-492.
3. Rosendal M, Olde Hartman TC, Aamland A, et al. "Medically unexplained" symptoms and symptom disorders in primary care: Prognosis-based recognition and classification. *BMC Family Practice*. 2017;18(1). doi: 10.1186/s12875-017-0592-6.
4. Houwen J, Lucassen PLBJ, Stappers HW, Assendelft WJJ, Dulmen Sv, Olde Hartman TC. Improving GP communication in consultations on medically unexplained symptoms: A qualitative interview study with patients in primary care. *British Journal of General Practice*. 2017;67(663):e716-e723. doi: 10.3399/bjgp17X692537.
5. Pescosolido BA. Patient trajectories. In: Cockerham WC, Dingham R, Quah SR, eds. *The wiley blackwell encyclopedia of health, illness, behavior, and society.* 1st ed. Hoboken, U.S.A.: Wiley Online Library; 2014:1770-1777.
6. Barsky AJ, Peekna HM, Borus JF. Somatic symptom reporting in women and men. *Journal of General Internal Medicine*. 2001;16(4):266-275. doi: 10.1046/j.1525-1497.2001.00229.x.
7. Ballering AV, Wardenaar KJ, olde Hartman TC, Rosmalen JGM. Female sex and femininity independently associate with common somatic symptom trajectories. *Psychological Medicine*. 2022;52(11):2144-2154. doi: 10.1017/S0033291720004043.
8. Ballering AV, Bonvanie IJ, Olde Hartman TC, Monden R, Rosmalen JGM. Gender and sex independently associate with common somatic symptoms and lifetime prevalence of chronic disease. *Social Science & Medicine*. 2020;253:112968. doi: 10.1016/j.socscimed.2020.112968.
9. Jackson JL, Chamberlin J, Kroenke K. Gender and symptoms in primary care practices. *Psychosomatics*. 2003;44(5):359-366. doi: 10.1176/appi.psy.44.5.359.
10. Green MH. *The trotula: A medieval compendium of women's medicine.* 1st ed. Philadelphia, U.S.A.: University of Pennsylvania Press; 2001:1-301.
11. Lippi D, Bianucci R, Donell S. Gender medicine: Its historical roots. *Postgraduate medical journal*. 2020;96:480-486. doi: 10.1136/postgradmedj-2019-137452.
12. Adair MJ. Plato's view of the 'wandering uterus'. *The Classical Journal*. 1996;91(2):153-163.
13. Tasca C, Rapetti M, Carta MG, Fadda B. Women and hysteria in the history of mental health. *Clinical practice and epidemiology in mental health*. 2012;8:110-119. doi: 10.2174/1745017901208010110.
14. Bem SL. The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*. 1974;42(2):155-162. doi: 10.1037/h0036215.
15. Horstmann S, Schmechel C, Palm K, Oertelt-Prigione S, Bolte G. The operationalisation of sex and gender in quantitative Healthâ€"Related research: A scoping review. *International Journal of Environmental Research and Public Health*. 2022;19(12). doi: 10.3390/ijerph19127493.
16. Choi N, Fuqua DR, Newman JL. The bem sex-role inventory: Continuing theoretical problems. *Educational and Psychological Measurement*. 2008;68(5):881-900. doi: 10.1177/0013164408315267.
17. Lagro-Janssen T. Sex, gender, and health: Developments in medical research. In: Kuhlmann E, Annandale E, eds. *The palgrave handbook of gender and healthcare*. London: Palgrave MacMillan; 2012:439-454. rg/10.1057/9781137295408_27.

18. Becher E, Oertelt-Prigione S. History and development of sex- and gender sensitive medicine (SGSM). *International Review of Neurobiology*. 2022;164:1-25. doi: 10.1016/bs.irn.2022.06.008.
19. Lippa R, Connelly S. Gender diagnosticity: A new bayesian approach to gender-related individual differences. *Journal of Personality and Social Psychology*. 1990;59(5):1051-1065. doi: 10.1037/0022-3514.59.5.1051.
20. Oertelt-Prigione S. Putting gender into sex- and gender-sensitive medicine. *eClinicalMedicine*. 2020;20:100305. doi: 10.1016/j.eclinm.2020.100305.
21. Tannenbaum C, Ellis RP, Eyssel F, Zou J, Schiebinger L. Sex and gender analysis improves science and engineering. *Nature*. 2019;575(7781):137-146. doi: 10.1038/s41586-019-1657-6.
22. Mauvais-Jarvis F, Bairey Merz N, Barnes PJ, et al. Sex and gender: Modifiers of health, disease, and medicine. *The Lancet*. 2020;396(10250):565-582. doi: 10.1016/S0140-6736(20)31561-0.
23. van den Hurk L, Hiltner S, Oertelt-Prigione S. Operationalization and reporting practices in manuscripts addressing gender differences in biomedical research: A cross-sectional bibliographical study. *International Journal of Environmental Research and Public Health*. 2022;19(21):14299. doi: 10.3390/ijerph192114299.
24. Ballering AV, Burke SM, Maeckelberghe ELM, Rosmalen JGM. How to ensure inclusivity in large-scale general population cohort studies? lessons learned with regard to including and assessing sex, gender, and sexual orientation. *Archives of Sexual Behavior*. 2023(52):2163–2172. doi: 10.1007/s10508-023-02600-y.
25. Johnson JL, Repta R. Sex and gender: Beyond the binaries. In: Oliffe JL, Greaves L, eds. *Designing and conducting gender, sex, & health research*. SAGE Publications, Inc; 2012:17-38. 10.4135/9781452230610.
26. Johnson JL, Greaves L, Repta R. Better science with sex and gender: Facilitating the use of a sex and gender-based analysis in health research. *International Journal for Equity in Health*. 2009;8(1):14. doi: 10.1186/1475-9276-8-14.
27. Göttgens I, Modderkolk L, Jansen C, Darweesh SKL, Bloem BR, Oertelt-Prigione S. The salience of gender in the illness experiences and care preferences of people with parkinson's disease. *Social Science & Medicine*. 2023;320:115757. doi: 10.1016/j.socscimed.2023.115757.
28. MacLean A, Sweeting H, Hunt K. 'Rules' for boys,'guidelines' for girls: Gender differences in symptom reporting during childhood and adolescence. *Social Science & Medicine*. 2010;70(4):597-604. doi: 10.1016/j.socscimed.2009.10.042.
29. Ballering AV, Oertelt-Prigione S, LCRI, Olde Hartman TC, Rosmalen JGM. Sex and gender-related differences in COVID-19 diagnoses and SARS-CoV-2 testing practices during the first wave of the pandemic: The dutch lifelines COVID-19 cohort study. *Journal of Women's Health (Larchmt.)*. 2021;30(12):1686-1692. doi: 10.1089/jwh.2021.0226.
30. Butler J. Performative acts and gender constitution: An essay in phenomenology and feminist theory. *Theatre Journal*. 1988;40(4):519-531. doi: 10.2307/3207893.
31. Smith PM, Koehoorn M. Measuring gender when you don't have a gender measure: Constructing a gender index using survey data. *International Journal for Equity in Health*. 2016;74(6):459-468. doi: 10.1186/s12939-016-0370-4.
32. Pelletier R, Ditto BF, Pilote L. A composite measure of gender and its association with risk factors in patients with premature acute coronary syndrome. *Psychosomatic medicine*. 2015;77(5):517-526. doi: 10.1097/PSY.0000000000000186.
33. West C, Zimmerman DH. Doing gender. *Gender and Society*. 1987;1(2):125-151.
34. Hankivsky O. Women's health, men's health, and gender and health: Implications of intersectionality. *Soc Sci Med*. 2012;74(11):1712-1720. doi: 10.1016/j.socscimed.2011.11.029.

35. Alswat KA. Gender disparities in osteoporosis. *Journal of Clinical Medicine Research*. 2017;9(5):382-387. doi: 10.14740/jocmr2970w.
36. Noh JW, Park H, Kim M, Kwon YD. Gender differences and socioeconomic factors related to osteoporosis: A cross-sectional analysis of nationally representative data. *Journal of Women's Health*. 2018;27(2). doi: 10.1089/jwh.2016.6244.
37. Corbin JM, Strauss AL. *Unending work and care: Managing chronic illness at home*. 1st ed. San Francisco, CA: Jossey-Bass Publishers; 1988.
38. Hannigan B, Evans N. Critical junctures in health and social care: Service user experiences, work and system connections. *Social Theory & Health*. 2013;11(4):428-444. doi: 10.1057/sth.2013.16.
39. Jutel A. Medically unexplained symptoms and the disease label. *Social Theory & Health*. 2010;8(3):229-245. doi: 10.1057/sth.2009.21.
40. Wikman A, Marklund S, Alexanderson K. Illness, disease, and sickness absence: An empirical test of differences between concepts of ill health. *Journal of Epidemiology and Community Health*. 2005;59(6):450-454. doi: 10.1136/jech.2004.025346.
41. Chiong W. Diagnosing and defining disease. *JAMA*. 2001;285(1):89-90. doi: 10.1001/jama.285.1.89-JMS0103-2-1.
42. Ballering AV, Olde Hartman TC, Verheij R, Rosmalen JGM. Sex and gender differences in primary care help-seeking for common somatic symptoms: A longitudinal study. *Scandinavian Journal of Primary Health Care*. 2023;41(2):132-139. doi: 10.1080/02813432.2023.2191653.
43. Plug I, van Dulmen S, Stommel W, Olde Hartman TC, Das E. Physicians' and patients' interruptions in clinical practice: A quantitative analysis. *Annals of Family Medicine*. 2022;20(5):423-429. doi: 10.1370/afm.2846.

PART 1

Etiology of common somatic symptoms

CHAPTER 2

Ballering, A.V., Ori, A.P.S., & Rosmalen, J.G.M. (2021). The association of sex, age and FKBP5 genotype with common somatic symptoms: A replication study in the lifelines cohort study. *Journal of Psychosomatic Research*, *147*, 110510.

Abstract

Objective: Our aim was to replicate a recent study that reported an association between the rs9470080 CC-genotype and common somatic symptoms in women, but not in men. Additionally, we quantified the genetic contribution to phenotypic variation in common somatic symptom levels.

Methods: We used data from the Lifelines Cohort Study, including 28,299 participants (60.0% female; 44.2% CC-genotype; mean age 42.9 (14.2) years). Common somatic symptoms were measured with the SCL-90 SOM subscale. To assess the association between the rs9470080 genotype and SCL-90 SOM scores we applied similar analyses as the original study, including independent *t*-tests, two-way ANOVAs and a mixed ANOVA. To estimate the proportion of phenotypic variance in SCL-90 SOM scores explained by single nucleotide polymorphisms (SNPs), we used a genomic-relatedness-based restricted maximum-likelihood method.

Results: We could not replicate the original study's findings. We found no association between the rs9470080 genotype and common somatic symptom levels in either female or male participants ($F(1, 8775)=1.07$, $p=0.30$ and $F(1,13,903)=0.01$, $p=0.93$, respectively). Genome-wide heritability analyses show that 12.1% ($p=2.1e-08$) of the phenotypic variance in common somatic symptom levels in Lifelines can be explained by SNPs. The genetic contribution to common somatic symptom levels was higher in male participants (SNP-h^2=20.5%; $p=9.1e-08$) than in female participants (SNP-h^2=12.0%, $p=2.8e-05$).

Conclusion: Our findings of significant SNP-h^2 and the sex-specific differences herein, does warrant further sex-stratified research of individual genetic variants associated with common somatic symptoms. Preferably, further research should be performed within the analytic framework of a genome-wide association study.

Introduction

Sex is increasingly recognized as a pivotal concept in health research.[1,2] In many diseases, including autoimmune disorders and cardiovascular disease, studies found sex differences in prevalence and presentation.[3,4] Similarly, sex differences are present in the distribution and presentation of common somatic symptoms: women are found to report more numerous, more intense and more frequent somatic symptoms than men.[5-7] Female sex also associates with a worse prognosis of common somatic symptoms.[8]

On the one hand, sex differences in the prevalence and longevity of common somatic symptoms are thought to associate with biological attributes, such as differences between male and female anatomy, hormones and genes.[7-9] On the other hand, gender, the psychosocial equivalent of biological sex encompassing the embodiment of different roles, behaviors, identities and relationships of men and women prescribed by social norms, also affects the prevalence of common somatic symptoms. Previous studies show that a sex-by-gender role interaction associates with common somatic symptoms.[7,8] This may point toward a gene-by-environment (GxE) interaction associating with common somatic symptom levels.

A recent study reported that rs9470080 CC-genotype, a single nucleotide polymorphism (SNP) in the FK506-binding protein 5 gene (*FKBP5*), associated with higher levels of common somatic symptoms in female participants, but not in male participants. However, the cohort was small (N=1,060), as was the effect size.[10] *FKBP5* is involved in the functioning of the HPA-axis.[11] The authors argue that the observed sex difference in common somatic symptoms may be attributed to sex differences in activity of the central nervous system. The authors, however, express the need for further studies to assess the importance of *FKBP5*, and for replication of their study in an independent cohort. We replicated the original study in the Lifelines Cohort Study. We furthermore quantified the contribution of SNPs to phenotypic variation in common somatic symptom levels to establish genome-wide SNP-heritability (SNP-h^2) of common somatic symptom levels in Lifelines.

Methods

Study design

In this study we used data from the Lifelines Cohort Study. Lifelines is a multi-disciplinary prospective population-based cohort study examining in a unique three-generation design the health and health-related behaviors of 167,729 persons living in the North

of The Netherlands. It employs a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioral, physical and psychological factors which contribute to the health and disease of the general population, with a special focus on multi-morbidity and complex genetics. Extensive information on the cohort and recruitment procedures is provided elsewhere.[12] Lifelines is performed according to the principles of the Declaration of Helsinki and is approved by the Medical Ethical Committee of the University Medical Center Groningen (number: 2007/152). For our analyses, we used data from three subsequent measurements. Participants were followed up on average after 13 (min-max: 10-93) and 25 (min-max: 22-92) months.

Variables

We assessed common somatic symptoms in the past week by the 12-item ordinal Symptom CheckList-90 Somatization subscale (SCL-90 SOM). The SCL-90 SOM refers to how much bother or distress participants experienced in the past 7 days due to somatic symptoms. Symptoms included, but were not limited to headache, dizziness and nausea.[8] The scale is recommended for large-scale studies and has sufficient measurement invariance over time.[13,14] Individual mean SCL-90 SOM scores were calculated for each timepoint.

Lifelines' genotyping, imputation procedures and quality control of genotype data were performed using standard protocols.[15] Participants' age and sex assigned at birth were derived from the municipal databases. As participants' sex refers to sex assigned at birth, we refer to participants as male or female.

Statistical analyses

For the replication analyses, we implemented a similar analyses pipeline as the original study. That is, we extracted rs9470080 (minor allele frequency [MAF]=0.33) and dichotomized the genetic variant into a CC- and CT/TT-genotype group. Similarly, we only included participants aged between 18 and 60 years, and grouped age by ≤49 years and ≥50 years.

As per the original study we conducted independent T-tests to cross-sectionally analyze the differences in common somatic symptom levels between age groups and sex. To assess the association between rs9470080 genotype and common somatic symptom levels, we applied a two-way ANOVA, adjusted for age groups. The cross-sectional analyses were based on 28,299 genotyped participants without missing data on included variables (60.0% female; 44.2% CC-genotype; 78.9% aged ≤49 years). To assess longitudinal data, we conducted a mixed ANOVA. Longitudinal analyses were based on 22,684 genotyped participants (61.3% female; 44.3% CC-genotype; 76.7% aged ≤49 years). We adhered to a two-sided α-value of 0.05. Aforementioned analyses were conducted in IBM SPSS v. 25.

Additionally, we conducted a genome-wide heritability analysis in 13,548 unrelated individuals, adjusted for sex and age. As SCL-90 SOM scores were non-normally distributed, we calculated average SCL-90 SOM scores across three time points per individual. Subsequently, we applied rank-based inverse normal transformation to generate a normally distributed trait. Using a genomic-relatedness-based restricted maximum-likelihood method we then estimated the proportion of phenotypic variance in SCL-90 SOM scores that is explained by all common SNPs (i.e. SNP-heritability/ SNP-h^2). In this analysis, we did not test for associations of individual SNPs. SNPs with (a) >5% missing data; (b) deviating from the Hardy-Weinberg equilibrium (p<1e-06) and; (c) with a MAF<0.01 were excluded from the genome-wide heritability analysis. We restricted our analysis to unrelated individuals (i.e. individuals with <5% degree of relatedness). The analysis was performed using PLINK and GCTA software.[16-18] In compliance with the SAGER guidelines, we report our results stratified by sex.[19]

Results

We found statistically significant differences in SCL-90 SOM scores between age groups, however these differences were negligible (***Table 1***). In contrast, we found no significant difference in common somatic symptom levels between rs9470080 genotype groups in male or female participants.

The two-way ANOVA showed that the main effect of female sex, adjusted for age groups, on SCL-90 SOM scores was statistically significant: F(1, 28294)=563.0, p<0.001. The effects of neither the rs9470080 variant (F(1,28294)=0.40, p<0.53), nor the sex-by-genotype interaction term (F(1,28294)=1.05, p=0.31) were statistically significant. These results indicate that in Lifelines, rs9470080 genotype did not associate with SCL-90 SOM scores, nor did this association differ in strength between female and male participants in the cross-sectional analyses.

As ***Table 2*** shows, we observed no association between rs9470080 genotype and SCL-90 SOM scores across three time points in neither male or female participants (F(1-8775)=1.07, p=0.30 and F(1-13,903)=0.01, p=0.93, respectively). Female sex significantly associated with SCL-90 SOM scores (F(1-22,679)=589.7, p<0.001). The sex-by-genotype interaction term was not statistically significant (F(1-22,697)=0.48, p=0.49), indicating that no significant sex difference in the association between rs9470080 genotype and common somatic symptom levels was present in longitudinal analyses.

Table 1. Independent T-tests to assess mean differences in SCL-90 SOM scores between age groups and rs9470080 genotype groups.

			Age			
	N	**Mean (SD)**	**DF**	**t**	***p*-value**	**Mean difference (95% CI)**
Male participants						
≤49 years	8,898	1.15 (0.22)	11,318	-3.59	<0.001	-0.02 (-0.03 - -0.01)
≥50 years	2,422	1.17 (0.24)				
Female participants						
≤49 years	13,431	1.22 (0.26)	16,977	-7.50	<0.001	-0.04 (-0.05 - -0.02)
≥50 years	3,548	1.26 (0.31)				
			rs9470080 genotype			
	N	**Mean (SD)**	**DF**	**t**	***p*-value**	**Mean difference (95% CI)**
Male participants						
CC-genotype	4,971	1.16 (0.23)	11,318	0.21	0.84	0.00 (-0.01 - 0.01)
CT/TT-genotype	6,349	1.15 (0.23)				
Female participants						
CC-genotype	7,531	1.23 (0.27)	16,977	-1.26	0.21	-0.01 (-0.01 - 0.00)
CT/TT-genotype	9,448	1.23 (0.27)				

We also assessed to which degree genetic factors contributed to phenotypic variation in SCL-90 SOM scores by means of genome-wide heritability analyses. In contrast to the analysis of a single genetic variant, this analysis estimated the variance in SCL-90 SOM scores explained by all common genetic variants in Lifelines. We observed a significant SNP-h^2 of 12.1% (N=13,548, p=2.1e-08) in SCL-90 SOM scores across three time points in unrelated individuals. Importantly, we observed a significantly higher genetic contribution in male participants (SNP-h^2= 20.5%, p=9.1e-08) than in female participants (SNP-h^2=12.0%, p=2.8e-05).

Table 2. Mixed ANOVA with common somatic symptom levels assessed by mean SCL-90 SOM score as an outcome.

		Between subject effects			
	DF	**Mean Square**	**F**	***p*-value**	**Effect size**
Male participants (N=8,778)					
rs9470080 (CT/TT)	1	0.07	1.07	0.30	0.00
Age (≥50 years)	1	0.96	14.1	<0.001	0.00
Female participants (N=13,906)					
rs9470080 (CT/TT)	1	0.01	0.01	0.93	0.00
Age (≥50 years)	1	4.31	47.3	<0.001	0.03
Total (N=22,684)					
rs9470080 (CT/TT)	1	0.05	0.65	0.42	0.00
Age(≥50 years)	1	4.98	60.6	<0.001	0.03
Sex (female)	1	48.3	589.7	<0.001	0.03
Sex (female) by rs9470080 (CT/TT)	1	0.04	0.48	0.49	0.00
	Within subject effects				
	DF	**Mean Square**	**F**	***p*-value**	**Effect size**
Male participants (N=8,778)					
Somatic symptoms	1.35	141.5	4494.5	<0.001	0.34
Somatic symptoms by rs9470080 (CT/TT)	1.35	0.01	0.39	0.596	0.00
Somatic symptoms by age (≥50 years)	1.35	0.15	4.71	0.020	0.01
Female participants (N=13,906)					
Somatic symptoms	1.36	321.1	8966.7	<0.001	0.39
Somatic symptoms by rs9470080 (CT/TT)	1.36	0.01	0.31	0.65	0.00
Somatic symptoms by age (≥50 years)	1.36	0.01	0.25	0.70	0.00
Total (N=22,684)					
Somatic symptoms	1.36	421.3	12341.2	<0.001	0.35
Somatic symptoms by rs9470080 (CT/TT)	1.36	0.01	0.39	0.60	0.00
Somatic symptoms by age (≥50 years)	1.36	0.09	2.51	0.10	0.00
Somatic symptoms by sex (female)	1.36	5.08	148.7	<0.001	0.01
Somatic symptoms by age (≥50 years) by sex (female)	1.36	0.01	0.31	0.65	0.00

As Mauchly's Test of Sphericity showed that the assumption of sphericity was violated, we included a Greenhouse-Geisser adjustment for the degrees of freedom (ε= 0.67 in males, ε= 0.68 in females, ε= 0.68 total).

Discussion

Despite the larger sample size of our study, we could not replicate the original study's findings. Cross-sectionally, we found significant, yet negligible differences in common somatic symptom levels between age groups in both male and female participants. In neither male nor female participants the symptom levels differed significantly between rs9470080 groups. Longitudinally, we could not corroborate the original study's finding of a significant association between rs9470080 genotype and common somatic symptom levels in female participants. Genome-wide heritability analyses show that 12.1% of the variance in common somatic symptom levels in Lifelines can be explained by common genetic variants, with a higher genetic contribution in male participants than in female participants. Sex differences in genetic contribution to disease have been previously reported, for example in depression.[20]

The discrepancy in results between the studies could be due to the differing times of follow-up or the different set of somatic symptoms that was assessed. The original study also included symptoms related to sleep, whereas our study focused solely on common somatic symptoms. Possibly, the association reported in the original study could also have been a chance finding as a result of a type I error.

A paucity of studies assessing common somatic symptoms in a genome-wide manner exists,[21] with merely one twin-study known to the authors that reported a 7-29% variation in somatic symptoms due to genotype.[22] It is thought that a polygenetic architecture underlies the experience of common somatic symptoms.[23] Therefore, it is highly likely that single SNPs do not associate with common somatic symptom levels, as shown in this study. This means that an interplay between a vast variety of genetic variants results in differing common somatic symptom levels. Therefore, we argue that the clinical relevance of single genetic variants is limited.[24] Nevertheless, our findings of significant SNP-h^2 and the sex-specific differences herein do warrant further investigations of individual genetic variants associated with common somatic symptoms. Such studies, however, should preferably be conducted within the analytic framework of a genome-wide association study and will likely require large sample sizes.

Lastly, although the original study focused on differences in a single SNP, we argue that sex differences in common somatic symptoms are more complex. Sex differences in the prevalence and persistence of common somatic symptoms are likely to be influenced by additional biological factors, including pain processing pathways, and environmental factors such as gender roles.[7,8] Notably, sex-related biological processes and gender may interact and have an effect on health.[25]

References

1. Regitz-Zagrosek V. Sex and gender differences in health. *EMBO reports*. 2012;13(7):596-603. doi: 10.1038/embor.2012.87.
2. Mauvais-Jarvis F, Bairey Merz N, Barnes PJ, et al. Sex and gender: Modifiers of health, disease, and medicine. *The Lancet*. 2020;396(10250):565-582. doi: 10.1016/S0140-6736(20)31561-0.
3. Spence JD, Pilote L. Importance of sex and gender in atherosclerosis and cardiovascular disease. *Atherosclerosis*. 2015;241(1):208-210. doi: 10.1016/j.atherosclerosis.2015.04.806.
4. Gold S, Willing A, Leypoldt F, Paul F, Friese M. Sex differences in autoimmune disorders of the central nervous system. *Seminars in Immunopathology*. 2019;41(2):177-188. doi: 10.1007/s00281-018-0723-8.
5. Barsky AJ, Peekna HM, Borus JF. Somatic symptom reporting in women and men. *Journal of General Internal Medicine*. 2001;16(4):266-275. doi: 10.1046/j.1525-1497.2001.00229.x.
6. Tomenson B, Essau C, Jacobi F, et al. Total somatic symptom score as a predictor of health outcome in somatic symptom disorders. *The British Journal of Psychiatry*. 2013;203(5):373-380. doi: 10.1192/bjp.bp.112.114405.
7. Ballering AV, Bonvanie IJ, Olde Hartman TC, Monden R, Rosmalen JGM. Gender and sex independently associate with common somatic symptoms and lifetime prevalence of chronic disease. *Social Science & Medicine*. 2020;253:112968. doi: 10.1016/j.socscimed.2020.112968.
8. Ballering AV, Wardenaar KJ, olde Hartman TC, Rosmalen JGM. Female sex and femininity independently associate with common somatic symptom trajectories. *Psychological Medicine*. 2022;52(11):2144-2154. doi: 10.1017/S0033291720004043.
9. Johnson JL, Greaves L, Repta R. Better science with sex and gender: Facilitating the use of a sex and gender-based analysis in health research. *International Journal for Equity in Health*. 2009;8(1):14. doi: 10.1186/1475-9276-8-14.
10. Sannes A, Christensen JO, Nielsen MB, Gjerstad J. The influence of age, gender and the FKBP5 genotype on subjective health complaints in the norwegian working population. *Journal of Psychosomatic Research*. 2020;139:110264. doi: 10.1016/j.jpsychores.2020.110264.
11. Scharf SH, Liebl C, Binder EB, Schmidt MV, Müller MB. Expression and regulation of the Fkbp5 gene in the adult mouse brain. *PLoS ONE*. 2011;6(2):e16883. doi: 10.1371/journal.pone.0016883.
12. Klijs B, Scholtens S, Mandemakers JJ, Snieder H, Stolk RP, Smidt N. Representativeness of the LifeLines cohort study. *PLoS ONE*. 2015;10(9):e0137203. doi: 10.1371/journal.pone.0137203.
13. Zijlema WL, Stolk RP, Löwe B, Rief W, White PD, Rosmalen JGM. How to assess common somatic symptoms in large-scale studies: A systematic review of questionnaires. *Journal of Psychosomatic Research*. 2013;74(6):459-68. doi: 10.1016/j.jpsychores.2013.03.093.
14. Rytilä-Manninen M, Fröjd S, Haravuori H, et al. Psychometric properties of the symptom checklist-90 in adolescent psychiatric inpatients and age- and gender-matched community youth. *Child and Adolescent Psychiatry and Mental Health*. 2016;10(23):23. doi: 10.1186/s13034-016-0111-x.
15. Neustaeter A, Nolte I, Snieder H, Jansonius NM. Genetic pre-screening for glaucoma in population-based epidemiology: Protocol for a double-blind prospective screening study within lifelines (EyeLife). *BMC Ophthalmology*. 2021;21(1). doi: 10.1186/s12886-020-01771-9.
16. Lee S, Wray N, Goddard M, Visscher P. Estimating missing heritability for disease from genome-wide association studies. *American journal of human genetics*. 2011;88(3):294-305. doi: 10.1016/j.ajhg.2011.02.002.

17. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A tool for genome-wide complex trait analysis. *American Journal of Human Genetics*. 2011;88(1):76-82. doi: 10.1016/j.ajhg.2010.11.011.
18. Purcell S, Neale B, Todd-Brown K, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*. 2007;81(3):559-575. doi: 10.1086/519795.
19. Heidari S, Babor TF, De Castro P, Tort S, Curno M. Sex and gender equity in research: Rationale for the SAGER guidelines and recommended use. *Research Integrity and Peer Review*. 2016;1:2. doi: 10.1186/s41073-016-0007-6.
20. Kendler KS, Gardner CO, Neale MC, Prescott CA. Genetic risk factors for major depression in men and women: Similar or different heritabilities and same or partly distinct genes? *Psychological Medicine*. 2001;31(4):605-616. doi: 10.1017/S0033291701003907.
21. Holliday KL, Macfarlane GJ, Nicholl BI, Creed F, Thomson W, McBeth J. Genetic variation in neuroendocrine genes associates with somatic symptoms in the general population: Results from the EPIFUND study. *Journal of Psychosomatic Research*. 2010;68(5):469-474. doi: 10.1016/j.jpsychores.2010.01.024.
22. Kato K, Sullivan PF, Pedersen NL. Latent class analysis of functional somatic symptoms in a population-based sample of twins. *Journal of Psychosomatic Research*. 2010;68(5):447-453. doi: 10.1016/j.jpsychores.2010.01.010.
23. Duncan LE, Ostacher M, Ballon J. How genome-wide association studies (GWAS) made traditional candidate gene studies obsolete. *Neuropsychopharmacology*. 2019;44(9):1518-1523. doi: 10.1038/s41386-019-0389-5.
24. Lewis CM, Vassos E. Polygenic risk scores: From research tools to clinical instruments. *Genome Medicine*. 2020;12(44). doi: 10.1186/s13073-020-00742-5.
25. Alabas OA, Tashani OA, Tabasam G, Johnson MI. Gender role affects experimental pain responses: A systematic review with meta-analysis. *European Journal of Pain*. 2012;16(9):1211-1223. doi: 10.1002/j.1532-2149.2012.00121.x.

CHAPTER 3

Hogendoorn, E., Ballering, A.V., van Dijk, M.W., Rosmalen, J.G.M, & Burke, S.M. (2023). Discordance between Adolescents and Parents in Functional Somatic Symptom Reports: Sex Differences and Future Symptom Prevalence. *Journal of Youth and Adolescence*, 52(10), 2182-2195.

Abstract

Functional somatic symptoms, i.e., physical complaints that cannot be sufficiently explained by an objectifiable biomedical abnormality, become increasingly more prevalent in girls than in boys during adolescence. Both parents and adolescents report more functional somatic symptoms in girls, but their reports correspond only limitedly. It remains unknown whether parent-adolescent discordance contributes to the higher symptom prevalence in girls. This study investigated parent-adolescent discordance in reported functional somatic symptoms throughout adolescence, examined the longitudinal association of parent-adolescent discordance with symptom prevalence in early adulthood and focused on sex differences in these processes. Participants included 2229 adolescents (50.7% female) from four assessments (age 11 to 22 years) of the TRAILS population cohort. Parents and adolescents reported significantly more symptoms in girls than in boys during adolescence. Variance analyses showed that throughout adolescence, parents reported fewer symptoms than girls self-reported and more than boys self-reported. Regression analyses using standardized difference scores showed that lower parent-report than self-report was positively associated with symptom prevalence in early adulthood. Polynomial regression analyses revealed no significant interaction between parent-reported and adolescent self-reported symptoms. Associations did not differ between boys and girls. The findings show that lower parent-reported than self-reported symptoms predict future symptom prevalence in both sexes, but this discordance was more observed in girls. The higher functional somatic symptom prevalence in girls might be partly explained by parental underestimation of symptoms.

Introduction

During adolescence, there is a growing difference between boys' and girls' reporting of functional somatic symptoms (i.e., physical complaints that cannot be sufficiently explained by a detectable biomedical abnormality), with girls reporting more symptoms than boys.[1] Gendered parenting with regard to symptoms in adolescence may be a factor contributing to this sex difference. Gaining more insight into discordance in parent-reported and adolescent boys' and girls' self-reported functional somatic symptoms is crucial to understand the higher symptom prevalence in girls and targeting parents for interventions. However, the majority of studies only include either parent-report or self-report rather than a combination, and lack follow-ups into adulthood. Thus, the current four-wave longitudinal study aimed to investigate parent-adolescent discordance in reporting of functional somatic symptoms in boys and girls over the course of adolescence and its association with symptom prevalence in early adulthood.

Functional somatic symptoms are defined as physical symptoms, for example headache, back pain or tiredness that cannot be entirely attributed to a detectable biomedical abnormality after adequate diagnostic research and history taking.[2] Experiencing functional somatic symptoms from time to time is normal and these complaints usually spontaneously disappear. However, 4% of adolescents experience persistent functional somatic symptoms.[3] Persistent symptoms in adolescence are associated with physical impairment, deteriorated school functioning, and social withdrawal.[4] Girls tend to consistently report more functional somatic symptoms than boys.[5,6] The sex difference in symptom pre-valence is already present in childhood and tends to increase during adolescence.[7] This could be due to biological dissimilarities, as during the adolescent period many physical changes take place, which possibly influence somatic symptom proneness.[8]

Aside from differences in biological vulnerability for somatic symptoms, the sex difference in symptom pre-valence may relate to psychosocial differences. Psychosocial differences between boys and girls can be described in terms of gender. Gender is an umbrella term entailing the embodiment of different identities, roles, behaviors and relationships of men and women prescribed by societal norms in a given time and society, whereas sex refers to biological characteristics, including hormones and anatomy of male and female bodies.[9] Research has shown that in adults, gender associates with somatic symptoms independently of sex, which is possibly due to the adherence to normative gender roles.[10] The traditional masculine gender roles include stoicism, high pain tolerance and not showing weakness, whereas feminine

gender roles allow for vulnerability and expression of pain.[11] During adolescence, most boys and girls behave increasingly according to their socially-prescribed gender roles.[12] In addition, adolescents generally become more independent of their parents or caregivers (henceforth referred to as parents) and spend more time with pre-dominantly same-sex peers, which encourages adherence to gender roles.[13,14] A qualitative study revealed that adolescents are aware of gender role expectations regarding somatic symptoms and that they feel pressured to adhere to these, especially when among peers.[15] Thus, these processes of gendered socialization in adolescence may be related to symptom reporting in boys and girls.

Gender role patterns regarding symptoms may be transmitted from parents to their children through family upbringing. Social learning of illness behavior begins in childhood. From an early age onwards, children learn to interpret physical sensations, give meaning to them and respond to them by observing and communicating with their caregivers.[16] As such, parents' management of their child's symptoms may influence the child's interpretation, communication and management of future symptoms. Consistent with gender role expectations, parents encourage more independence and control of emotions ("being tough") in sons regarding symptoms, while they behave more protectively towards daughters, and encourage daughters to share their feelings and symptom experiences.[17,18] Any reaction to a symptom is preceded by an assessment of that symptom. The way parents assess their child's symptoms differs for boys and girls, with meta-analytic evidence showing that parents report more functional somatic symptoms in girls than in boys, possibly reflecting parental beliefs about gender roles.[19]

However, the previous meta-analysis and subsequent studies did not include adolescents' self-reported functional somatic symptoms.[3,19] Yet, prior studies have found that parent-reports and child self-reports on somatic symptoms correspond only to a limited extend.[20,21] In addition, it has been shown that parent-child discordance in psychopathology predicts several clinical features, such as emotional and behavioral problems and social competence.[22,23] Combining both parent and child perspectives is clinically relevant and useful to inform preventive, diagnostic and treatment strategies for functional somatic symptoms.[24] Previous studies have identified high parent-reported functional somatic symptoms as a risk factor for persistence of symptoms, but dis-regarded sex differences and included follow-ups to mid-adolescence instead of adulthood.[3] Parent-child discordance in functional somatic symptoms may reflect gendered parenting. This is indicated by the finding that parents behave differently towards sons and daughters when they experience somatic symptoms, and may thus perceive symptoms differently.[18] It has also been suggested that parents are

more likely to report symptoms in their children that are congruent with gender expectations.[25] Parent-adolescent discordance in symptom reporting could thus be informative in studying sex differences in functional somatic symptoms. It would be highly valuable to gain more insight into the course of (discordance between) parent- and self-reported functional somatic symptoms throughout adolescence, its association with symptom pre-valence in adulthood, and sex differences herein.

The current study

There is a paucity of studies that combine parent and adolescent perspectives and examine sex differences in the longitudinal course and associations of parent-adolescent discordance in symptom reporting. Therefore, it remains unclear whether parent-adolescent discordance contributes to the higher functional somatic symptom prevalence in girls. This study sought to address the gaps in the existing literature by taking into account the sex of the adolescent, including both parent-reported and adolescent self-reported functional somatic symptoms and discordance herein, and by adopting a longitudinal approach with follow-ups into early adulthood. This study examined differences in functional somatic symptoms in boys and girls over the course of adolescence, studying parent-report and self-report (aim 1). Furthermore, this study investigated if parent-adolescent discordance changes over time in adolescence and differs between boys and girls (aim 2). Lastly, longitudinal associations were investigated between parent-adolescent discordance and symptom prevalence in early adulthood, and sex differences herein (aim 3). This was studied in a large population-based cohort using four assessment waves in adolescence and early adulthood. Based on literature on gendered parenting, it was hypothesized that parent-adolescent discordance in reported functional somatic symptoms is larger in adolescent girls than in boys, with parents perceiving more symptoms in their daughters than their sons (hypothesis 1). Furthermore, this study hypothesized that the course of parent-adolescent discordance is different for boys and girls (hypothesis 2). Lastly, based on literature indicating that parent-child discordance in psychopathology contributes to future poor outcomes, it was expected that parent-adolescent discordance predicts symptom prevalence in early adulthood (hypothesis 3a), and that increased parental reporting of functional somatic symptoms in girls, compared to adolescent self-report, contributes to the higher symptom prevalence in girls in early adulthood (hypothesis 3b).

Method

Sample and procedure

This study is part of the Tracking Adolescents' Individual Lives (TRAILS) study. TRAILS is an ongoing prospective cohort study that investigates mental health and social development from pre-adolescence onwards. TRAILS-participants lived in one of the three northern provinces of the Netherlands at the time of recruitment and were intended to represent the general Dutch preadolescent population. Participants were recruited through primary schools. Primary schools that participated in TRAILS were comparable to other primary schools in the Netherlands with regard to the proportion of children with a low socioeconomic background. Detailed information about recruitment and sample characteristics has been reported elsewhere.[26] Topics of previous TRAILS studies include the use of reports of multiple informants, and have pointed out the informative nature of discrepancies in these reports, highlighting the importance of studying longitudinal associations of discordance among informants.[27]

Participants enrolled in the TRAILS study at age 10–12 years. Measurement waves have been taking place bi- or triennially. In the current study, data from T1 (mean age 11.1 years, 51% female), T2 (mean age 13.6 years, 51% female), T3 (mean age 16.3 years, 52% female) and T5 (mean age 22.3 years, 53% female) were used from the complete sample (n = 2230, of which one parent-child dyad had their data deleted upon parental request, resulting in a sample size of 2229). The Dutch Central Committee on Research Involving Human Subjects (CCMO) granted ethical approval for the TRAILS study (#NL38237.042.11). Parents provided written informed consent at T1. At T2, T3 and T5, written informed consent was also obtained from the TRAILS-participant.

Measures

Main variables of the study involved self-reported and parent-reported functional somatic symptoms. SES, pubertal status and gender non-contentedness were included as covariates as previous studies indicated these may be relevant to sex differences in parent-adolescent discordance of somatic symptoms.[1,28-30]

Self-reported functional somatic symptoms

Self-reported functional somatic symptoms were measured using the Somatic Complaints subscale of the Youth Self Report (YSR) at T1, T2 and T3.[31] At T5, the Adult Self Report (ASR) was used, which was appropriate for the age of the participants at that time.[32] This subscale contains items that refer to somatic complaints without a known medical cause or without obvious reason. The TRAILS-participant indicated to what extent each complaint had applied to him/her in the prior six months. Answers

were rated on a three-point Likert scale (0 = not at all true; 1 = sometimes true; 2 = often true). Two items, 'eye problems' and 'skin problems', were excluded since previous TRAILS studies reported low factor loadings, indicating that these two items did not represent the underlying construct of functional somatic symptoms well in the TRAILS cohort.[1,3] Moreover, three items of the ASR (heart pounding, numbness, and trouble sleeping) were excluded to ensure consistency with the YSR. The remaining seven items included dizziness, overtiredness, aches/pains, headache, nausea, stomach pain, and vomiting. Both the YSR and ASR show adequate reliability, validity and measurement invariance.[33-35] Internal consistency (measured by Cronbach's alpha) of the seven items used in our study was good (T1: $\alpha = 0.76$; T2: $\alpha = 0.77$; T3: $\alpha = 0.75$; T5: $\alpha = 0.71$).

Parent-reported functional somatic symptoms

Parent-reported functional somatic symptoms were assessed at T1, T2 and T3 using the Somatic Complaints subscale of the Child Behavior Checklist (CBCL).[36] The CBCL corresponds to the YSR and ASR presented in the form of parent-report. The parent of the TRAILS-participant stated to what extent each complaint had applied to their child in the prior six months on a three-point Likert scale with the same scoring categories. The item 'obstipation' was excluded from the parent-reported data since it is not part of the YSR/ASR, leaving only the corresponding items. Psychometric properties are good and the CBCL has been validated in numerous populations.[37,38] Internal consistency of the seven items used in our study was good (T1: $\alpha = 0.71$; T2: $\alpha = 0.72$; T3: $\alpha = 0.73$).

Sex

Sex of the TRAILS-participant was dichotomously assessed at T1 using self-report (male/female).

Gender non-contentedness

In this study, gender non-contentedness refers to any expressed or felt desire to be of the opposite gender or sex. Gender non-contentedness was measured at T1, T2, T3 and T5 with the item 'I wish to be of the opposite sex' of the YSR or ASR. Adolescents experiencing gender non-contentedness might show sex-incongruent gender role behaviors with regard to their symptoms, which also might influence parents' assessment of their child's symptoms. The TRAILS-participant indicated to what extent this statement had applied to him/her in the prior six months on a 3-point Likert scale (0 = not at all true; 1 = sometimes true; 2 = often true). Gender non-contentedness was defined as a score equaling or exceeding 1.

Socioeconomic status

Socioeconomic status (SES) was assessed at T1 by calculating the average of the z-scores of the following indicators: educational and occupational level of each parent, and household income. Z-scores were calculated based on the International Standard Classification of Occupations.[39]

Pubertal status

Pubertal status was measured at T1, T2 and T3 using the Tanner Scale of Pubertal Status.[40,41] The Tanner Scale includes five stages, with a higher stage referring to a later stage of development. At T1, the parent of the TRAILS-participant was surveyed about two physical characteristics of their child. Using schematic drawings, they reported on genital development and pubic hair for boys, and breast development and pubic hair for girls. For each characteristic, the parent rated which of the five Tanner stages was most applicable to the TRAILS-participant. At T2 and T3, pubertal status was assessed by a self-reported five-item questionnaire. The characteristics for boys comprised growth spurt, skin changes, body hair, voice-change, and facial-hair growth. The characteristics for girls included growth spurt, skin changes, body hair, breast development, and menarche. Answers were rated on a 4-point Likert-scale (0 = not yet started; 1 = barely started; 2 = definitely started; 3 = seems complete), except for menarche, which was assessed dichotomously (0 = no, 1 = yes).

Statistical analyses

This study was preregistered prior to analysis of the data (https://osf.io/cbrqa). Characteristics of the study sample are presented per assessment wave. According to the SAGER guidelines, results were stratified by sex if applicable.[42]

First, to examine if parents and adolescents perceive functional somatic symptoms differently in boys than in girls over the course of adolescence (aim 1), correlations between parent-reported and self-reported functional somatic symptoms at T1, T2 and T3 were calculated for boys and girls separately. Independent T-tests were performed to assess whether parent-reported and self-reported functional somatic symptoms at T1, T2 and T3 differed statistically significantly between boys and girls. Subsequently, standardized difference scores were calculated by subtracting standardized parent-reported symptom scores from standardized self-reported symptom scores at T1, T2 and T3. Then, to assess whether the standardized difference between parent-reported and adolescent-reported symptoms differed per sex of the adolescent, ANCOVA tests were performed at T1, T2 and T3 with sex as fixed effect and parent-adolescent discordance as dependent variable. SES, pubertal status and gender non-contentedness were included as covariates.

Second, sex differences in changes in parent-adolescent discordance over the course of adolescence were tested (aim 2) using a mixed model ANOVA. Sex was included as between-subjects factor, time as within-subjects factor, and parent-adolescent discordance as dependent variable. The repeated measures of T1, T2 and T3 were used.

Lastly, it was examined whether differences in parent-reports and adolescent self-reports of functional somatic symptoms provide information in the prediction of symptom prevalence in early adulthood (aim 3). Two statistical approaches were applied to incorporate informant discordance in the prediction of later symptom prevalence. According to an earlier study, using standardized difference scores when predicting health outcomes requires caution, because it may yield inaccurate results due to unequal variability and different bivariate associations in reports of different informants.[43] Examination of informant interaction terms in a polynomial regression framework is therefore recommended. Yet, another study compared the use of standardized difference scores and polynomial regression, and concluded that both approaches can be used complementary to each other to provide more nuanced and comprehensive results regarding discordance in informant reporting.[44]

An advantage of combining both approaches is that two slightly distinct hypotheses are tested, i.e., whether a mere difference between informant reports associates with the outcome (using standardized difference scores), and whether the association between the report of one informant and the outcome varies as a function of the report of the other informant (using polynomial regression). Therefore, in the current study, it was examined if parent-adolescent discordance at T1 through T3 predicted functional somatic symptom prevalence at T5 using both approaches. In the first part, using the method of standardized difference scores, it was initially planned to regress symptom prevalence at T5 on parent-adolescent discordance at T1–T3 in a multi-level model, to account for dependency of residual errors. The model did, however, not converge when doing so. Therefore, deviating from the preregistration, the same regression was conducted but now in a linear regression model.

First, the assumption of independent residuals of the repeated measures of functional somatic symptoms was checked using visual inspection of scatterplots and the Durbin-Watson statistic. Parent-adolescent discordance at T1–T3, expressed in the standardized difference score of parent-reported and self-reported functional somatic symptoms at T1–T3, was included as independent variable and self-reported functional somatic symptoms at T5 as dependent variable. Adolescent sex, SES,

pubertal status and gender non-contentedness were included as covariates. Parent-adolescent discordance at T1–T3 by adolescent sex was entered as interaction term, to test whether the association differed for boys and girls. All continuous predictors were standardized before including them in the model.

In the second part, polynomial regression analyses were conducted, again using symptom prevalence at T5 as dependent variable and adolescent sex, SES, pubertal status and gender non-contentedness as covariates. Now, parent-reported functional somatic symptoms at T1-T3 and self-reported functional somatic symptoms at T1-T3 were included as separate independent variables (instead of a combination resulting in one variable reflecting parent-adolescent discordance). First, the assumption of independent residuals was checked. Second, the effects of covariates were entered (block 1), followed by the main effects[a], the quadratic main effects[b] and linear interaction terms[c] (block 2). Third, cubic main effects[d] and quadratic interaction terms[e] were entered (block 3), but dropped again if model fit did not significantly improve and none of the additional interaction terms was significant. In post-hoc polynomial regression analyses, it was assessed if the longitudinal associations between parent-reported functional somatic symptoms at T1-T3 and self-reported functional somatic symptoms at T5 differed for boys and girls, by adding an interaction between parent-reported functional somatic symptoms at T1-T3 and sex. Corresponding higher-order terms were entered in the same block-wise manner as in the main analyses. Analyses were conducted in SPSS version 28. Unstandardized regression coefficients with 95% confidence intervals and *p*-values are reported. An α-level of .005 was applied to correct for multiple testing.

Multiple imputation of missing data

The percentages of missing data for self-reported functional somatic symptoms were 5.2% at T1, 6.2% at T2, 12.3% at T3, and 16.2% at T5. The percentages of missing data for parent-reported functional somatic symptoms were 11.3% at T1, 14.3% at T2, and 19.5% at T3. Missingness is unlikely to be completely random. Therefore, multiple imputation was applied in the longitudinal analyses to minimalize the risk of bias. Five data sets were generated using the Series Mean Imputation procedure in SPSS. All data sets were analyzed in an identical way, whereafter the results were pooled using Rubin's rules.[45]

Results

Descriptive results

Table 1 shows the characteristics of the study sample stratified by assessment wave and sex. Both parent-reported and self-reported functional somatic symptoms were significantly higher in girls than in boys at all three waves (T1–T3). As shown by sex-stratified Pearson correlations, parent-reported and self-reported functional somatic symptoms were significantly positively correlated at weak to moderate strength at all three waves (see ***Appendix A***).

Table 1. Descriptives of the study sample

		T1	T2	T3	T5
Boys	n (%)[a]	1098 (49.26)	1054 (49.07)	867 (47.69)	843 (47.33)
	Age – M (SD)	11.13 (0.56)	13.57 (0.52)	16.28 (0.71)	22.34 (0.63)
	SES – M (SD)	-0.07 (0.82)	n/a	n/a	n/a
	Self-reported FSS – M (SD)[b]	3.05 (2.41)[d]	2.17 (2.26)[e]	1.59 (1.85)[f]	1.15 (1.50)
	Parent-reported FSS – M (SD)[c]	1.29 (1.73)[g]	1.08 (1.57)[h]	0.78 (1.29)[i]	n/a
	Pubertal status – M (SD)	1.71 (0.58)	2.61 (1.08)	2.72 (0.79)	n/a
	Gender non-contentedness – n (%)[a]	125 (11.8)	45 (4.4)	28 (3.6)	12 (1.8)
Girls	n (%)[a]	1131 (50.74)	1094 (50.93)	951 (52.31)	938 (52.67)
	Age – M (SD)	11.09 (0.55)	13.57 (0.54)	16.28 (0.71)	22.25 (0.67)
	SES – M (SD)	-0.03 (0.78)	n/a	n/a	n/a
	Self-reported FSS – M (SD)[b]	3.45 (2.47)[d]	3.22 (2.55)[e]	3.09 (2.51)[f]	2.61 (2.34)
	Parent-reported FSS – M (SD)[c]	1.54 (1.81)[g]	1.52 (1.83)[h]	1.62 (2.02)[i]	n/a
	Pubertal status – M (SD)	2.01 (0.87)	3.74 (0.93)	2.84 (0.53)	n/a
	Gender non-contentedness – n (%)[a]	143 (12.9)	88 (8.2)	55 (6.2)	31 (3.7)

Note. FSS = functional somatic symptoms; n/a = not applicable. [a] = percentage based on total sample without missing data; [b] = sum score of seven included items of the Somatic Complaints subscale of the YSR (T1-T3) or ASR (T5); [c] = sum score of seven included items of the Somatic Complaints subscale of the CBCL; [d]: Independent T-test: $t(2112) = 3.814$, $p < .001$; [e]: Independent T-test: $t(2012) = 9.785$, $p < .001$; [f]: Independent T-test: $t(1592) = 13.520$, $p < .001$; [g]: Independent T-test: $t(1976) = 3.093$, $p = .002$; [h]: Independent T-test: $t(1839) = 35.503$, $p < .001$; [i]: Independent T-test: $t(1462) = 9.308$, $p < .001$.

Discordance in parent- and self-reported functional somatic symptoms in adolescence

Figure 1 shows the variability of standardized parent-child discordance for boys and girls at each wave in adolescence. For both boys and girls, the medians at all waves are close to zero. Variability appears to be greater for girls.

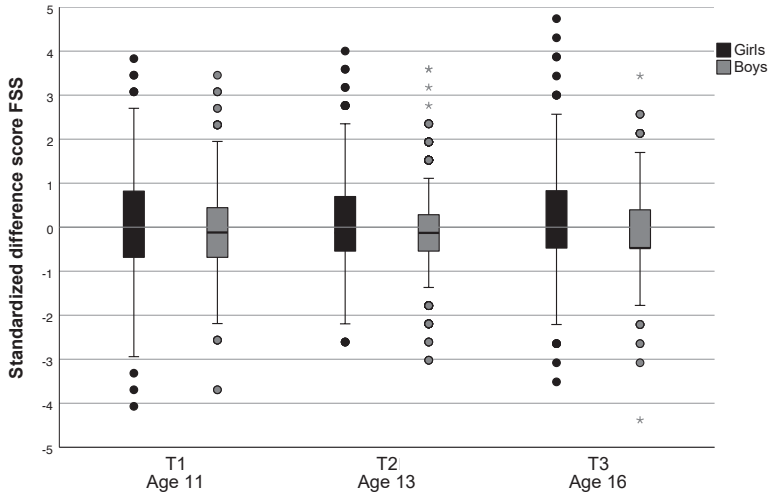


Figure 1: Boxplots of parent-adolescent discordance per measurement wave expressed in standardized difference scores, clustered by sex.

Note: Standardized difference scores were calculated by subtracting standardized parent-reported functional somatic symptoms from standardized self-reported functional somatic symptoms.

ANCOVAs were performed per assessment wave to test if parent-adolescent discordance in functional somatic symptoms differed per sex of the adolescent. At T1, no significant association was found between sex and standardized parent-adolescent discordance after adjusting for the effect of SES, gender non-contentedness and pubertal status: F(1, 1815) = 0.413, $p = 0.520$. At T2 and T3, standardized parent-adolescent discordance in functional somatic symptoms differed significantly between boys and girls after adjusting for the effect of SES, gender non-contentedness and pubertal status (T2: $F(1, 1707) = 8.312$, $p = 0.004$; T3: $F(1, 1292) = 31.359$, $p < 0.001$). The standardized difference was negative in boys, indicating that parents reported more symptoms than boys themselves. In girls, the standardized difference was positive, indicating that parents reported fewer symptoms than girls themselves. A mixed ANOVA was performed to test if parent-adolescent discordance changed over the course of adolescence and if this differed per sex. No significant main effect of time on parent-adolescent discordance was found ($F(2, 2154) = 0.925$, $p = 0.397$). However, both a significant main effect of sex on parent-adolescent discordance ($F(1, 1077) = 25.779$, $p < 0.001$) and a time-by-sex interaction effect on parent-adolescent discordance ($F(2, 2154) = 7.757$, $p < 0.001$) were found. ***Figure 2*** visualizes how standardized parent-adolescent discordance developed over time for boys and girls, with parents reporting slightly less symptoms than girls themselves, and slightly more symptoms than boys themselves over time.

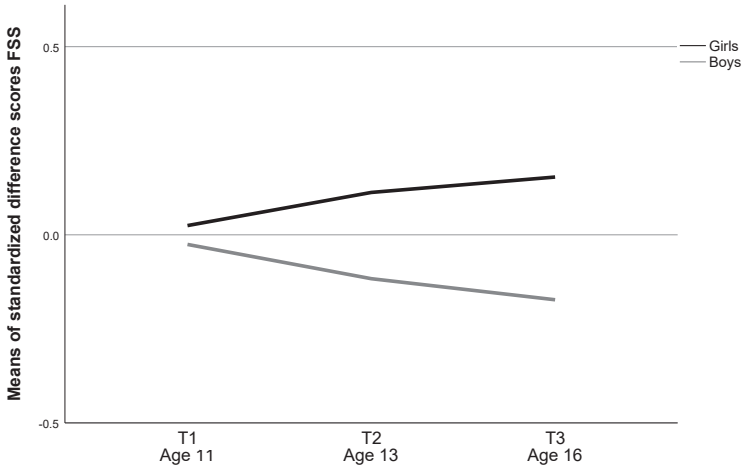


Figure 2. Means of the standardized difference between parent-reported and self-reported functional somatic symptoms over time, clustered by sex.

Note. Standardized difference scores were calculated by subtracting standardized parent-reported functional somatic symptoms from standardized self-reported functional somatic symptoms.

Longitudinal associations between discordance and functional somatic symptoms in early adulthood

Linear regression analyses were conducted to test if parent-adolescent discordance at T1–T3, expressed in standardized difference scores, predicted symptom prevalence at T5. The assumption of independent residuals was met, indicating that there was no dependence between the residuals of parent-adolescent discordance across the waves. Zero order correlations between parent-reported FSS at T1–T3 and self-reported FSS at T5, and self-reported FSS at T1–T3 and self-reported FSS at T5, respectively, were in the same direction and range (see ***Appendix B***). This shows similar bivariate correlations between the reports of parent-report and self-report and the outcome.

Table 2 shows that parent-adolescent discordance at T1–T3 significantly associated with self-reported functional somatic symptoms at T5, after adjusting for sex, SES, pubertal status and gender non-contentedness ($b = 0.234$, 95% CI [0.134, 0.334], $p < 0.001$). This indicates that higher positive parent-adolescent discordance (i.e., lower parent-reported functional somatic symptoms than self-reported functional somatic symptoms) was associated with higher self-reported functional somatic symptoms at T5. The interaction between parent-adolescent discordance and sex was not significant ($b = -0.057$, 95% CI [−0.193, −0.079], $p = 0.396$). This indicates that the strength of the association between parent-adolescent discordance at T1–T3 and self-reported functional somatic symptoms at T5 does not differ between boys and girls.

Table 2. Linear regression analyses: longitudinal associations between parent-adolescent discordance in functional somatic symptoms at T1-T3 and self-reported functional somatic symptoms at T5

Predictor	β (95% CI)	p
Parent-adolescent discordance T1-T3	0.234 (0.134 – 0.334)	<.001
Adolescent sex	-0.922 (-1.060 – -0.784)	<.001
Parent-adolescent discordance * Adolescent sex	-0.057 (-0.193 – 0.079)	.396
SES	-0.311 (-0.386 – -0.237)	<.001
Gender non-contentedness	0.084 (-0.105 – 0.272)	.383
Pubertal status	0.056 (0.004 – 0.108)	.035

Polynomial regression analyses were conducted to test if the association between parent-reported functional somatic symptoms at T1–T3 and symptom prevalence at T5 differed by adolescents' self-reported functional somatic symptoms at T1–T3. The assumption of independent residuals was met, indicating that there was no dependence between the residuals of parent-reported functional somatic symptoms across the waves and, likewise, the residuals of self-reported functional somatic symptoms across the waves. Model fit did not significantly improve by entering cubic main effects and quadratic interaction terms (block 3), and none of the higher-order interactions were significant. Quadratic main effects were also non-significant. This indicates that polynomial terms did not fit the data better than linear terms.

Table 3 shows that both parent-reported functional somatic symptoms at T1–T3 ($b = 0.314$, 95% CI [0.236, 0.392], $p < 0.001$) and self-reported functional somatic symptoms at T1–T3 ($b = 0.426$, 95% CI [0.363, 0.489], $p < 0.001$) were significant predictors of self-reported functional somatic symptoms at T5. However, the two-way interaction term between self-reported functional somatic symptoms at T1–T3 and parent-reported functional somatic symptoms at T1–T3 was not significant ($b = -0.011$, 95% CI [−0.076, 0.054], $p = 0.729$). This indicates that the strength of the association between parent-reported functional somatic symptoms at T1–T3 and self-reported functional somatic symptoms at T5 does not differ by self-reported functional somatic symptoms at T1–T3 (i.e., the association is similar for high and low self-reported functional somatic symptoms at T1–T3). The three-way interaction term between self-reported functional somatic symptoms T1–T3, parent-reported functional somatic symptoms T1–T3, and adolescent sex was also not statistically significant ($b = 0.047$, 95% CI [−0.043, 0.136], $p = 0.306$), indicating that the strength of the association between parent-reported functional somatic symptoms at T1–T3 and self-reported functional somatic symptoms at T5 does not differ by self-reported functional somatic symptoms at T1–T3 between boys and girls.

Table 3. Polynomial regression analyses: longitudinal associations between parent- and self-reported functional somatic symptoms at T1-T3 and self-reported functional somatic symptoms at T5

Predictor	β (95% CI)	p
Adolescent sex	-0.735 (-0.856 – -0.614)	<.001
SES	-0.200 (-0.273 – -0.127)	<.001
Gender non-contentedness	-0.030 (-0.210 – 0.150)	.741
Pubertal status	0.052 (0.002 – 0.102)	.041
Parent-reported FSS T1-T3	0.314 (0.236 – 0.392	<.001
Self-reported FSS T1-T3	0.426 (0.363 – 0.489)	<.001
Parent-reported FSS T1-T3^2	-0.006 (-0.050 – 0.039)	.793
Self-reported FSS T1-T3^2	0.034 (-0.010 – 0.078)	.126
Parent-reported FSS T1-T3 * Self-reported FSS T1-T3	-0.011 (-0.076 – 0.054)	.729
Parent-reported FSS T1-T3 * Self-reported FSS T1-T3 * Adolescent sex	0.047 (-0.043 – 0.136)	.306

Note. FSS = functional somatic symptoms; Cubic main effects and quadratic interaction terms were entered in block 3, but were dropped from the model again since model fit did not significantly improve and none of the additional interaction terms were significant. These higher-order predictors were therefore omitted from the table.

Post-hoc analyses revealed no significant parent-reported functional somatic symptoms at T1-T3 by sex interaction (see ***Appendix C***), indicating no difference between boys and girls in strength of the association between parent-reported functional somatic symptoms at T1-T3 and self-reported functional somatic symptoms at T5.

Discussion

Although extensive research has shown that functional somatic symptoms are more prevalent in girls than in boys, a gap exists in the literature regarding the role of discordance in parent-reported and adolescent self-reported symptoms in the sex difference in symptom prevalence. To address this gap, this study investigated sex differences in parent-adolescent discordance in reported functional somatic symptoms and its longitudinal association with future symptom prevalence using data from four waves of a large population-based cohort.

The results showed that parent-reported and self-reported functional somatic symptoms were significantly higher in girls than in boys throughout adolescence. However, when comparing parent-report with self-report, parents reported slightly less symptoms than girls self-reported. The course of parent-adolescent discordance

over adolescence differed between boys and girls, with parents increasingly reporting more symptoms in girls than girls themselves over time, and vice versa in boys. Furthermore, using standardized difference scores, it was found that lower parent-reported than self-reported symptoms contributed to symptom prevalence in early adulthood. Using polynomial regression, it was found that interactions between parent-reported and self-reported symptoms did not associate with early adulthood symptom prevalence. No sex differences were detected in the associations with early adulthood symptom prevalence.

The weak to moderate correlations between parent-reported and self-reported functional somatic symptoms that we found are consistent with previous studies assessing informant discordance, which reported correlations in the range of 0.15–0.40.[20,21,46,47] An explanation for low correspondence could lie in the relatively low observability of somatic symptoms, which may be even more the case with functional somatic than in symptoms that are part of a biomedical condition. The finding that higher levels of functional somatic symptoms are reported in girls, both by themselves and by their parents, also concurs with previous work.[6,19,48] The sex difference that the current study found in parent-child discordance in adolescence, however, contrasts with previous studies on parent-adolescent discordance in somatic symptoms reports, which showed no moderation effect of sex.[49,50] Those studies had, in contrast to the current study, cross-sectional designs and included younger participants, which may explain the different findings.

Contrary to the hypothesis, parental underestimation of symptoms was observed in girls and overestimation in boys. This finding contrasts with gender role literature on femininity, stating that girls are more susceptible to pain and more open in expressing their complaints as it is more socially accepted for women and girls to express pain.[15] An explanation might be, albeit speculative, that girls do behave increasingly according to the feminine gender role, thus more openly expressing symptoms, and boys vice versa. However, parents may not perceive this behavioral change, resulting in parental underestimation of symptoms in girls and overestimation in boys. Alternatively, transmission of symptom-related gender roles may not be captured well by studying parent-adolescent discordance, as symptom management comprises more than merely the estimation of symptom prevalence. Alternatively, transmission of symptom-related gender roles may not be well captured by studying parent-adolescent discordance, as symptom management comprises more than merely the estimation of symptom prevalence. Possibly, it would be better captured in measures reflecting social learning, such as illness behavior modelling or parental responses to child symptoms.

This is the first large epidemiological study that assessed longitudinal associations between parent-adolescent discordance and future symptom prevalence. Confirming the hypothesis, it was found that parent-adolescent discordance, expressed in standardized difference scores, contributed to future symptoms. This finding is similar to previous studies that focused on parent-child discordance in psychopathology and showed longitudinal associations with later adverse psychological outcomes, such as anxiety and depressive symptoms.[23,51] Notably, using polynomial regression analysis, the interaction between parent-reported and self-reported symptoms in adolescence was not significantly associated with early adulthood symptom prevalence. The results show that parent-reported symptoms in adolescence associated independently with early adulthood symptom prevalence, but the strength of this association was unaffected by self-reported symptoms in adolescence. Yet, the existence of discordance between parent-report and self-report, measured using standardized difference scores, was associated with early adulthood symptom prevalence. Contrary to the hypothesis, however, the discordance contributing to future symptoms concerned parental underestimation (parents reporting fewer symptoms than the adolescent), rather than parental overestimation (parents reporting more symptoms than the adolescent). Possibly, parents failing to identify symptoms in their children could lead to the child feeling misunderstood or overlooked, which may be harmful for their health and well-being.[23] Moreover, adolescents with burdening functional somatic symptoms may not receive professional treatment in time because their parents do not see the need to take their child to a health care professional, while early interventions could prevent persistence and exacerbation of symptoms.[52] In this way, parental underestimation of symptoms could contribute to a poor prognosis. It is worth noting that although parental underestimation of symptoms contributed to future symptom prevalence in both boys and girls, more underestimation was observed in girls. This possibly constitutes part of the explanation for the higher prevalence of functional somatic symptoms in girls.

The findings of this study, if replicated, may aid in developing preventive and treatment strategies for burdening functional somatic symptoms. Children of parents who report fewer symptoms than the children themselves, which in our study were predominantly girls, may be particularly at risk of a poor symptom prognosis. For clinicians working with adolescents with functional somatic symptoms, it is important to be aware of the limited correspondence between parent-reported and self-reported symptoms and of the risk posed by parental unawareness of symptoms. Important to mention is that high parent-reported FSS emerged as an independent risk factor for future symptom prevalence, concurring with previous studies.[3] This indicates

that preventive and treatment strategies may be desirable not only in the case of parental underestimation of symptoms but also in cases of high parental perception of symptoms (regardless of the adolescent's self-reported symptoms). These findings underline the importance of involving parents in the treatment of children and adolescents presenting with functional somatic symptoms, as has been indicated by previous research.[53]

Strengths and Limitations

This study has several strengths. We used data from a large population-based cohort, which enhances the generalizability of the results.[26] Moreover, we included data from four assessment waves covering the entire developmental period from adolescence into early adulthood. Another strength is the use of validated instruments to assess functional somatic symptoms by different informants, including adolescents. In addition, we included gender non-contentedness in our model. Gender, in addition to sex, may explain differences in the occurrence and trajectories of somatic symptoms[10,54] and may influence parent-reported and self-reported functional somatic symptoms. We have included this variable in an effort to capture an aspect of the complex interplay between sex, gender and somatic symptoms. Furthermore, we examined the longitudinal effect of informant discordance using both the approach of standardized difference scores and polynomial regression, allowing for more comprehensive, informative and nuanced results.[43,44]

Limitations of our study should be taken into account as well. First, parental sex could not be included in the study, as these data were unavailable. Parental sex may associate with levels of reported functional somatic symptoms in their child, as mothers and fathers may perceive symptoms in their child differently.[25] Furthermore, functional somatic symptoms were assessed using the YSR, ASR and CBCL. Even though these questionnaires specifically assess symptoms that occur without medical cause or obvious reason, we cannot be sure that the reported somatic symptoms are not part of an explained biomedical condition. However, regarding longitudinal associations of parent-adolescent discordance and functional somatic symptoms in early adulthood, it may not matter if the symptoms are medically explained or unexplained, as similar gendered socialization processes could take place. Finally, non-random attrition was present, with more drop-outs of male and low SES participants. However, no differences in internalizing problems (including functional somatic symptoms) were found between participants who continued to participate and those who dropped out.[26] Moreover, multiple imputation was used for the longitudinal analyses to handle missing data, thereby reducing the risk of bias.

Directions for Future Research

Possibly, parental underestimation of functional somatic symptoms in their daughters partly explains the higher symptom prevalence in girls. Other factors contributing to the increasing sex difference in functional somatic symptom prevalence are thought to include differences in biological features (e.g., hormones and pain regulatory systems), symptom labeling, puberty-related increase of depressive and anxiety symptoms in girls and incidence of sexual abuse.[55-58] Future research should focus on elucidating factors contributing to the growing sex difference in functional somatic symptom prevalence during adolescence. Studying sex differences in processes of social learning of illness behavior longitudinally, including assessments from an early age onwards, may further clarify the emergence of sex differences in functional somatic symptoms. Furthermore, considering the growing body of evidence showing that gender associates with health, future research should include measures of gender when studying sex differences in functional somatic symptoms in children and adolescents.[9,10,54,59,60] In addition, relations between parent-adolescent discordance and future functional somatic symptom prevalence may be different in clinical populations. Studying parent-adolescent discordance in relation to developmental trajectories of somatic symptoms in clinical samples could be informative in this respect.

Conclusion

There is a lack of studies combining parent-reported and adolescent boys' and girls' self-reported functional somatic symptoms to investigate the possible influence of parent-adolescent discordance on future symptom prevalence. The current study aimed to examine sex differences in the longitudinal course of parent-adolescent discordance and its association with symptom prevalence in early adulthood. The results of this study show that parental underestimation of functional somatic symptoms contributes to increased future symptom prevalence. In our study, this association was similar for boys and girls. However, over the course of adolescence, parental underestimation was more observed in girls than in boys. This might partly explain why girls are more prone to experiencing functional somatic symptoms than boys. The findings may aid in developing preventive and treatment strategies for burdening functional somatic symptoms. Interventions may be targeted specifically at parent-child dyads that differ greatly in their reporting. Future research should seek to elucidate factors contributing to the increasing sex difference in symptom prevalence in adolescence, which are presumably biopsychosocial.

Endnotes

[a]parent-reported functional somatic symptoms; self-reported functional somatic symptoms; [b]parent-reported functional somatic symptoms[2]; self-reported functional somatic symptoms[2]; [c]parent-reported functional somatic symptoms at T1-T3 by self-

reported functional somatic symptoms at T1-T3; parent-reported functional somatic symptoms at T1-T3 by self-reported functional somatic symptoms at T1-T3 by sex [d]parent-reported functional somatic symptoms^3; self-reported functional somatic symptoms^3; [e]parent-reported functional somatic symptoms^2 by self-reported functional somatic symptoms; parent-reported functional somatic symptoms by self-reported functional somatic symptoms^2; parent-reported functional somatic symptoms^2 by self-reported functional somatic symptoms by sex; parent-reported functional somatic symptoms by self-reported functional somatic symptoms^2 by sex

References

1. Janssens KAM, Rosmalen JGM, Ormel J, et al. Pubertal status predicts back pain, overtiredness, and dizziness in american and dutch adolescents. *Pediatrics*. 2011;128(3):553-559. doi: 10.1542/peds.2010-2364.
2. Beck JE. A developmental perspective on functional somatic symptoms. *Journal of pediatric psychology*. 2008;33(5):547-562. doi: 10.1093/jpepsy/jsm113.
3. Janssens KAM, Klis S, Kingma EM, Oldehinkel AJ, Rosmalen JGM. Predictors for persistence of functional somatic symptoms in adolescents. *The Journal of pediatrics*. 2014;164(4):900-905(e2). doi: rg/10.1016/j.jpeds.2013.12.003.
4. Youssef NN, Murphy TG, Langseder AL, Rosh JR. Quality of life for children with functional abdominal pain: A comparison study of patients' and parents' perceptions. *Pediatrics*. 2006;117(1):54-59. doi: 10.1542/peds.2005-0114.
5. Romero-Acosta K, Canals J, Hernández-Martínez C, Penelo E, Zolog TC, Domènech-Llaberia E. Age and gender differences of somatic symptoms in children and adolescents. *Journal of Mental Health*. 2013;22(1):33-41. doi: 10.3109/09638237.2012.734655.
6. van Geelen SM, Rydelius P, Hagquist C. Somatic symptoms and psychological concerns in a general adolescent population: Exploring the relevance of DSM-5 somatic symptom disorder. *Journal of Psychosomatic Research*. 2015;79(4):251-258. doi: 10.1016/j.jpsychores.2015.07.012.
7. LeResche L, Mancl LA, Drangsholt MT, Saunders K, Korff MV. Relationship of pain and symptoms to pubertal development in adolescents. *Pain*. 2005;118(1):201-209. doi: 10.1016/j.pain.2005.08.011.
8. Vogt Yuan AS. Gender differences in the relationship of puberty with adolescents' depressive symptoms: Do body perceptions matter? *Sex roles*. 2007;57(1-2):69-80. doi: 10.1007/s11199-007-9212-6.
9. Smith PM, Koehoorn M. Measuring gender when you don't have a gender measure: Constructing a gender index using survey data. *International Journal for Equity in Health*. 2016;74(6):459-468. doi: 10.1186/s12939-016-0370-4.
10. Ballering AV, Bonvanie IJ, Olde Hartman TC, Monden R, Rosmalen JGM. Gender and sex independently associate with common somatic symptoms and lifetime prevalence of chronic disease. *Social Science & Medicine*. 2020;253:112968. doi: 10.1016/j.socscimed.2020.112968.
11. Pool GJ, Schwegler AF, Theodore BR, Fuchs PN. Role of gender norms and group identification on hypothetical and experimental pain tolerance. *Pain*. 2007;129(1):122-129. doi: 10.1016/j.pain.2006.10.008.
12. Chaplin TM, Aldao A. Gender differences in emotion expression in children: A meta-analytic review. *Psychological bulletin*. 2013;139(4):735-765. doi: 10.1037/a0030737.
13. Lam CB, McHale SM, Crouter AC. Time with peers from middle childhood to late adolescence: Developmental course and adjustment correlates. *Child development*. 2014;85(4):1677-1693. doi: 10.1111/cdev.12235.
14. Witt SD. The influence of peers on children's socialization to gender roles. *Early Child Development and Care*. 2000;162(1):1-7. doi: 10.1080/0300443001620101.
15. MacLean A, Sweeting H, Hunt K. 'Rules' for boys,'guidelines' for girls: Gender differences in symptom reporting during childhood and adolescence. *Social Science & Medicine*. 2010;70(4):597-604. doi: 10.1016/j.socscimed.2009.10.042.

16. Levy RL, Whitehead WE, Von Korff MR, Feld AD. Intergenerational transmission of gastrointestinal illness behavior. *The American journal of gastroenterology*. 2000;95(2):451-456. doi: 10.1016/S0002-9270(99)00825-4.
17. Walker LS, Zeman JL. Parental response to child illness behavior. *Journal of pediatric psychology*. 1992;17(1):49-71. doi: 10.1093/jpepsy/17.1.49.
18. Clementi M, Faraji P, Poppert Cordts K, et al. Parent factors are associated with pain and activity limitations in youth with acute musculoskeletal pain: A cohort study. *The Clinical journal of pain*. 2019;35(3):222-228. doi: 10.1097/AJP.0000000000000668.
19. Duhig AM, Renk K, Epstein MK, Phares V. Interparental agreement on internalizing, externalizing, and total behavior problems: A meta-analysis. *Clinical psychology*. 2000;7(4):435-453. doi: 10.1093/clipsy.7.4.435.
20. De Los Reyes A, Augenstein TM, Wang M, et al. The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological bulletin*. 2015;141(4):858-900. doi: 10.1037/a0038498.
21. Hart SL, Hodgkinson SC, Belcher HME, Hyman C, Cooley-Strickland M. Somatic symptoms, peer and school stress, and family and community violence exposure among urban elementary school children. *Journal of Behavioral Medicine*. 2013;36(5):454-465. doi: 10.1007/s10865-012-9440-2.
22. Guion K, Mrug S, Windle M. Predictive value of informant discrepancies in reports of parenting: Relations to early adolescents' adjustment. *Journal of Abnormal Child Psychology*. 2009;37(1):17-30. doi: 10.1007/s10802-008-9253-5.
23. van de Looij-Jansen PM, Jansen W, de Wilde EJ, Donker MCH, Verhulst FC. Discrepancies between parent-child reports of internalizing problems among preadolescent children: Relationships with gender, ethnic background, and future internalizing problems. *The Journal of early adolescence*. 2011;31(3):443-462. doi: 10.1177/0272431610366243.
24. Kraemer HC, Measelle JR, Ablow JC, Essex MJ, Boyce WT, Kupfer DJ. A new approach to integrating data from multiple informants in psychiatric assessment and research: Mixing and matching contexts and perspectives. *The American journal of psychiatry*. 2003;160(9):1566-1577. doi: 10.1176/appi.ajp.160.9.1566.
25. Schroeder JF, Hood MM, Hughes HM. Inter-parent agreement on the syndrome scales of the child behavior checklist (CBCL): Correspondence and discrepancies. *Journal of Child and Family Studies*. 2010;19(5):646-653. doi: 10.1007/s10826-010-9352-0.
26. Oldehinkel AJ, Rosmalen JG, Buitelaar JK, et al. Cohort profile update: The tracking adolescents' individual lives survey (TRAILS). *International Journal of Epidemiology*. 2015;44(1):76-76n. doi: 10.1093/ije/dyu225.
27. Noordhof A, Oldehinkel AJ, Verhulst FC, Ormel J. Optimal use of multi-informant data on co-occurrence of internalizing and externalizing problems: The TRAILS study. *International journal of methods in psychiatric research*. 2008;17(3):174-183. doi: 10.1002/mpr.258.
28. Brix N, Ernst A, Lauridsen LLB, et al. Timing of puberty in boys and girls: A population-based study. *Paediatric and Perinatal Epidemiology*. 2019;33(1):70-78. doi: 10.1111/ppe.12507.
29. Hinz A, Ernst J, Glaesmer H, et al. Frequency of somatic symptoms in the general population: Normative values for the patient health questionnaire-15 (PHQ-15). *Journal of Psychosomatic Research*. 2017;96:27-31. doi: 10.1016/j.jpsychores.2016.12.017.
30. Roubinov DS, Boyce WT. Parenting and SES: Relative values or enduring principles? *Current opinion in psychology*. 2017;15:162-167. doi: 10.1016/j.copsyc.2017.03.001.
31. Achenbach TM, Rescorla LA. Manual for the ASEBA school-age forms & profiles. *University of Vermont, Research Center for Children, Youth, & Families*. 2001.

32. Achenbach TM, Rescorla LA. Manual for the ASEBA adult forms & profiles. *University of Vermont, Research Center for Children, Youth, & Families*. 2003.
33. Achenbach TM, Dumenci L, Rescorla LA. Ten-year comparisons of problems and competencies for national samples of youth: Self, parent, and teacher reports. *Journal of Emotional and Behavioral Disorders*. 2002;10(4):194-203. doi: 10.1177/10634266020100040101.
34. Fonseca-Pedrero E, Sierra-Baigrie S, Lemos-Giráldez S, Paino M, Muñiz J. Dimensional structure and measurement invariance of the youth self-report across gender and age. *The Journal of Adolescent Health*. 2012;50(2):148-153. doi: 10.1016/j.jadohealth.2011.05.011.
35. Najman JM, Heron MA, Hayatbakhsh MR, et al. Screening in early childhood for risk of later mental health problems: A longitudinal study. *Journal of psychiatric research*. 2008;42(8):694-700. doi: 10.1016/j.jpsychires.2007.08.002.
36. Achenbach TM. Manual for the child behavior checklist/ 4-18 and 1991 profile. *University of Vermont, Research Center for Children, Youth, & Familie*s. 1991.
37. Fombonne E. The use of questionnaires in child psychiatry research: Measuring their performance and choosing an optimal cut-off. *Journal of child psychology and psychiatry*. 1991;32(4):677-693. doi: 10.1111/j.1469-7610.1991.tb00343.x.
38. Verhulst FC, Akkerhuis GW, Althaus M. Mental health in dutch children: (I) a cross-cultural comparison. *Acta psychiatrica Scandinavica*. 1985;72(s323):1-108. doi: 10.1111/j.1600-0447.1985.tb10512.x.
39. Ganzeboom HBG, Treiman DJ. Internationally comparable measures of occupational status for the 1988 international standard classification of occupations. *Social science research*. 1996;25(3):201-239. doi: 10.1006/ssre.1996.0010.
40. Marshall WA, Tanner JM. Variations in pattern of pubertal changes in girls. *Archives of disease in childhood*. 1969;44(235):291-303. doi: 10.1136/adc.44.235.291.
41. Marshall WA, Tanner JM. Variations in pattern of pubertal changes in boys. *Obstetrical & gynecological survey*. 1970;45(239):13. doi: 10.1136/ADC.45.239.13.
42. Heidari S, Babor TF, De Castro P, Tort S, Curno M. Sex and gender equity in research: Rationale for the SAGER guidelines and recommended use. *Research Integrity and Peer Review*. 2016;1:2. doi: 10.1186/s41073-016-0007-6.
43. Laird RD, De Los Reyes A. Testing informant discrepancies as predictors of early adolescent psychopathology: Why difference scores cannot tell you what you want to know and how polynomial regression may. *Journal of Abnormal Child Psychology*. 2013;41(1):1-14. doi: 10.1007/s10802-012-9659-y.
44. Tackett JL, Herzhoff K, Reardon KW, Smack AJ, Kushner SC. The relevance of informant discrepancies for the assessment of adolescent personality pathology. *Clinical psychology*. 2013;20(4):378-392. doi: 10.1111/cpsp.12048.
45. Rubin DB. *Multiple imputation for nonresponse in surveys.* New York: Wiley; 1987.
46. Achenbach TM, McConaughy SH, Howell CT. Child/adolescent behavioral and emotional problems. *Psychological bulletin*. 1987;101(2):213-232. doi: 10.1037/0033-2909.101.2.213.
47. Sourander A, Helstelä L, Helenius H. Parent-adolescent agreement on emotional and behavioral problems. *Social Psychiatry and Psychiatric Epidemiology*. 1999;34(12):657-663. doi: 10.1007/s001270050189.
48. Janssens KAM, Zijlema WL, Joustra ML, Rosmalen JGM. Mood and anxiety disorders in chronic fatigue syndrome, fibromyalgia, and irritable bowel syndrome: Results from the LifeLines cohort study. *Psychosomatic Medicine*. 2015;77(4):449-457. doi: 10.1097/PSY.0000000000000161.

49. Kröner-Herwig B, Morris L, Heinrich M, Gassmann J, Vath N. Agreement of parents and children on characteristics of pediatric headache, other pains, somatic symptoms, and depressive symptoms in an epidemiologic study. *The Clinical journal of pain*. 2009;25(1):58-64. doi: 10.1097/AJP.0b013e31817fc62d.
50. Poulain T, Vogel M, Meigen C, Spielau U, Hiemisch A, Kiess W. Parent-child agreement in different domains of child behavior and health. *PLoS ONE*. 2020;15(4):e0231462. doi: 10.1371/journal.pone.0231462.
51. Ferdinand RF, van der Ende J, Verhulst FC. Parent-adolescent disagreement regarding psychopathology in adolescents from the general population as a risk factor for adverse outcome. *Journal of Abnormal Psychology*. 2004;113(2):198-206. doi: 10.1037/0021-843X.113.2.198.
52. Berezowski L, Ludwig L, Martin A, Löwe B, Shedden-Mora MC. Early psychological interventions for somatic symptom disorder and functional somatic syndromes: A systematic review and meta-analysis. *Psychosomatic medicine*. 2022;84(3):325-338. doi: 10.1097/PSY.0000000000001011.
53. Hulgaard D, Dehlholm-Lambertsen G, Rask CU. Family-based interventions for children and adolescents with functional somatic symptoms: A systematic review. *Journal of family therapy*. 2019;41(1):4-28. doi: 10.1111/1467-6427.12199.
54. Ballering AV, Wardenaar KJ, olde Hartman TC, Rosmalen JGM. Female sex and femininity independently associate with common somatic symptom trajectories. *Psychological Medicine*. 2022;52(11):2144-2154. doi: 10.1017/S0033291720004043.
55. Alloy LB, Hamilton JL, Hamlat EJ, Abramson LY. Pubertal development, emotion regulatory styles, and the emergence of sex differences in internalizing disorders and symptoms in adolescence. *Clinical psychological science*. 2016;4(5):867-881. doi: 10.1177/2167702616643008.
56. Meints SM, Edwards RR. Evaluating psychosocial contributions to chronic pain outcomes. *Progress in neuro-psychopharmacology & biological psychiatry*. 2018;87(Pt B):168-182. doi: 10.1016/j.pnpbp.2018.01.017.
57. Mendle J. Why puberty matters for psychopathology. *Child development perspectives*. 2014;8(4):218-222. doi: 10.1111/cdep.12092.
58. Barsky AJ, Peekna HM, Borus JF. Somatic symptom reporting in women and men. *Journal of General Internal Medicine*. 2001;16(4):266-275. doi: 10.1046/j.1525-1497.2001.00229.x.
59. Schiebinger L, Stefanick ML. Gender matters in biological research and medical practice. *Journal of the American College of Cardiology*. 2016;67(2):136-138. doi: 10.1016/j.jacc.2015.11.029.
60. Pelletier R, Ditto BF, Pilote L. A composite measure of gender and its association with risk factors in patients with premature acute coronary syndrome. *Psychosomatic medicine*. 2015;77(5):517-526. doi: 10.1097/PSY.0000000000000186.

Appendices

Appendix A Correlations between parent-reported and self-reported functional somatic symptoms

		Parent-reported FSS T1	Parent-reported FSS T2	Parent-reported FSS T3
Boys	Self-reported FSS T1	0.239; $p < .001$		
	Self-reported FSS T2		0.285; $p < .001$	
	Self-reported FSS T3			0.244; $p < .001$
Girls	Self-reported FSS T1	0.218; $p < .001$		
	Self-reported FSS T2		0.379; $p < .001$	
	Self-reported FSS T3			0.349; $p < .001$

Appendix B Correlations between parent-reported functional somatic symptoms in adolescence, self-reported functional somatic symptoms in adolescence and self-reported functional somatic symptoms in early adulthood

	Self-reported FSS T5
Self-reported FSS T1-T3	0.321; $p < .001$
Parent-reported FSS T1-T3	0.250; $p < .001$

Appendix C Post-hoc polynomial regression analyses: longitudinal associations between parent- and self-reported functional somatic symptoms at T1-T3 and self-reported functional somatic symptoms at T5**.**

Predictor	*b (95% CI)*	*p*
Adolescent sex	-0.742 (-0.868 – -0.617)	<.001
SES	-0.201 (-0.274 – -0.128)	<.001
Gender non-contentedness	-0.031 (-0.211 – 0.149)	.732
Pubertal status	0.052 (0.002 – 0.102)	.040
Parent-reported FSS T1-T3	0.353 (0.230 – 0.76	<.001
Self-reported FSS T1-T3	0.427 (0.364 – 0.490)	<.001
Parent-reported FSS T1-T3^2	-0.007 (-0.052 – 0.038)	.753
Self-reported FSS T1-T3^2	0.035 (-0.009 – 0.078)	.121
Parent-reported FSS T1-T3 * Self-reported FSS T1-T3	-0.024 (-0.088 – 0.040)	.461
Parent-reported FSS T1-T3 * Self-reported FSS T1-T3 * Adolescent sex	0.067 (-0.027 – 0.161)	.162
Parent-reported FSS T1-T3 * Adolescent sex	-0.081 (-0.233 – 0.070)	.273

Note. FSS = functional somatic symptoms; A two-way interaction between parent-reported FSS T1-T3 and adolescent sex was now included. Cubic main effects and quadratic interaction terms were entered in block 3, but were dropped again from the model since model fit did not significantly improve and none of the additional interaction terms were significant. These higher-order predictors were therefore omitted from the table.

PART 2

Prevalence and persistence of common somatic symptoms

CHAPTER 4

Ballering, A.V., Bonvanie, I.J., olde Hartman, T.C., Monden, R., & Rosmalen, J.G.M. (2020). Gender and sex independently associate with common somatic symptoms and lifetime prevalence of chronic disease. *Social Science & Medicine*, *253*, 112968.

Abstract

Sex and gender influence health differently. Associations between sex and health have been extensively studied, but gender (i.e. psychosocial sex) has been largely neglected, partly due to the absence of gender measures in cohort studies. Therefore, our objective was to test the unique associations of gender and sex with common somatic symptoms and chronic diseases, using a gender index created from existing cohort data. We applied LASSO logistic regression to identify, out of 153 unique variables, psychosocial variables that were predictive of sex (i.e. gender-related) in the Dutch Lifelines Cohort Study. These psychosocial variables covered gender roles and institutionalized gender. Using the estimated coefficients, gender indexes were calculated for each adult participant in the study (n=152,728; 58.5% female; mean age 44.6 (13.1) years). We applied multiple ordinal and logistic regression to test the unique associations of the gender index and sex, and their interactions, with common somatic symptoms assessed by the SCL-90 SOM and self-reported lifetime prevalence of chronic diseases, respectively. We found that in 10.1% of the participants the gender index was not in line with participants' sex: 12.5% of men and 8.4% of women showed a discrepancy between gender index and sex. Feminine gender characteristics are associated with increased common somatic symptoms and chronic diseases, especially in men. Female sex is associated with a higher common somatic symptom burden, but not with a higher prevalence of chronic diseases. The study shows that gender and sex uniquely impact health, and should be considered in epidemiological studies. Our methodology shows that consideration of gender measures in studies is necessary and feasible, based on data generally present in cohort studies.

Introduction

Sex and gender are increasingly recognized as essential aspects within health research.[1-3] Sex differences in the distribution and presentation of common somatic symptoms and medical conditions have been found, including in cardiovascular disease and depression, as well as in responses to treatments of these[4-6]. However, gender differences remain largely neglected in health research[7]. This is problematic, as evidence suggests that studying the roles of both sex and gender may reveal additional insights into their respective contribution in disease development, help-seeking behavior and response to treatment.[2,7,8]

To understand the importance of incorporating gender into health research, one should clearly distinguish between the concepts of sex and gender. Biological sex is defined as one's biological attributes, including physical features, chromosomes, gene expression, hormones and anatomy.[9] Yet, intersex variations exist in approximately 1.7% of the general population, challenging beliefs in absolute dimorphisms.[10] Gender, in contrast, can be seen as the psychosocial equivalent of biological sex: it encompasses the socially constructed roles, behaviors, identities and relationships of women, men and gender-diverse people in a given time and society.[8] In short, gender has a broader scope than sex and often refers to socially prescribed roles and behaviors, and experienced dimensions that relate to femininity and masculinity.[11] In general, gender roles, and the embodiment hereof, are not as static as one's biological sex usually is. They are subjected to ever-changing societal norms and institutions and may impact one's opportunities in life.[9]

Examples of differing associations of sex and gender roles with health can be found in multiple common medical conditions. For example, it is known that females develop osteoporosis more frequently due to biological differences in bone density and hormone levels.[12] Simultaneously, gender roles may encourage females (and feminine males) to restrict their food intake and perform sports not involving heavy weightlifting, further increasing their risk of developing osteoporosis.[13] Another example can be seen in healthcare seeking behavior and treatment allocation.[14] Masculine gender roles may affect help-seeking behavior and hamper adequate diagnosis and treatment, for example in depressive symptoms, as masculine gender roles are thought to be less expressive and to not openly acknowledge pain or impairments compared to feminine gender roles.[15,16] Disentangling associations of sex and gender with health may enhance our insights towards effective medicine in a given society.

Although the association of sex with health is often incorporated, the association of gender roles with health is seldom considered in research.[7,17] Therefore, little is known about whether sex and gender roles uniquely associate with symptoms and diagnoses. Similarly, whether these associations are present amongst a variety of health issues or merely specific symptoms, and whether associations differ between women and men, remains unknown. These gaps in knowledge may be attributable to difficulties in measuring gender. Existing gender measures, such as the Bem Sex Role Inventory (BSRI), Personal Attributes Questionnaire (PAQ) and gender diagnosticity measures have been extensively criticized.[18-20] These instruments measure gender via items that stereotype masculine and feminine characteristics, whilst gender roles are a broader concept largely dependent on time and place.[20-22] Possibly due to these difficulties, most epidemiological cohort studies do not measure gender in any way.

To the best of our knowledge, we present the first comprehensive analyses of the associations of sex and gender with health in a large epidemiological study: the general population cohort Lifelines. To this end, we constructed a gender index based on the existing data. We hypothesized that sex and the gender index will have a unique association with common somatic symptoms, as well as the lifetime prevalence of chronic disease. We also expected that female sex and feminine gender indexes will be associated with higher symptom levels and lifetime prevalence of chronic disease and that the associations of gender differ per sex.

Methods

Participants

This study was performed in Lifelines. Lifelines is a multi-disciplinary prospective population-based cohort study examining in a unique three-generation design the health and health-related behaviors of 167,729 persons living in the North of The Netherlands. It employs a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioral, physical and psychological factors which contribute to the health and disease of the general population, with a special focus on multi-morbidity and complex genetics. Extensive information on the cohort and recruitment is provided elsewhere.[23,24] The current study was based on the 152,728 adults included at baseline (***Table 1***). The Lifelines Cohort Study is performed according to the principles of the Declaration of Helsinki and in accordance with UMCG's research code. The Lifelines Cohort Study is approved by the Medical Ethical Committee of the University Medical Center Groningen, The Netherlands.

Table 1. Overview of demographic characteristics of the subsample on which the gender index is based and the complete adult Lifelines cohort at baseline.

Characteristic		Subsample Gender index (n=13,321)	Adult Lifelines Cohort (n=152,728)
Source into study (%)	Family	3,786 (28.3%)	49,264 (32.5%)
	GP	8,653 (64.4%)	81,533 (53.4%)
	Self-registered	956 (7.1%)	21,571 (14.1%)
Age in years (SD)		48.1 (11.4)	44.6 (13.1)
Sex (%)	Male	5,598 (41.8%)	63,388 (41.5%)
	Female	7,797 (58.2%)	89,340 (58.5%)
Median genderscore (IQR)	Male	0.05 (0.01-0.19)	0.06 (0.01-0.24)
	Female	0.97 (0.88-0.99)	0.96 (0.83-0.99)
Intersex conditions (%)		74 (0.55%)	1309 (0.86%)
Currently smoking (%)		3,113 (23.2%)	32,758 (21.4%)
Alcohol (%)	< 1 time a month	4,365 (32.6%)	31,195 (20.4%)
	1 to 3 times a month	3,827 (28.6%)	29,818 (19.5%)
	1 to 5 times a week	5,240 (39.1%)	75,141 (49.2%)
	6 to 7 times a week	1,448 (10.8%)	16,574 (10.9%)
Currently in a relationship (%)		11,663 (87.1%)	129,129 (84.5%)
Education (%)	No, primary or other education	768 (5.7%)	7380 (4.8%)
	Preparatory, vocational or junior secondary education	4,381 (32.7%)	41496 (27.2%)
	Senior secondary education or higher vocational education	3,885 (29.0%)	48741 (31.9%)
	University education	615 (4.6%)	9199 (6.0%)
Median SCL90SOM sumscore (IQR)	Male	1.17 (1.08-1.42)	1.17 (1.08-1.42)
	Female	1.33 (1.09-1.58)	1.33 (1.17-1.58)

Collected data

At baseline, participants completed questionnaires on topics including, but not limited to demographics, health, lifestyle and psychosocial aspects. Additionally, participants underwent physical examinations and biological samples, including DNA, were collected.[24] DNA material of 15,000 participants (9.8% of participating adults) was analyzed with 12-sample HumanCytoSNP-12 BeadChips. Quality control measures included i) exclusion of material with a call-rate <95%, ii) duplicate material,

iii) chromosome X heterozygosity was >0.005 (for municipally-registered males) or chromosome X heterozygosity was <0.1 (for municipally-registered females), and iv) material when sex chromosomes did not correspond with municipally-registered gender.

Subsample for calculation of the gender index
We conducted the analyses to calculate gender indexes on the subsample of participants from whom DNA was analyzed. We suspected that participants with an intersex condition or non-conform gender identity are more likely to have discrepancies between their psychosocial and biological sex, thus these were excluded from the analyses to compose the gender index.[25]

We defined intersex conditions as chromosomal variations of the sex-chromosomes (e.g. Triple-X syndrome); genetic mutation(s) resulting in hormonal disturbances relevant for sexual development (e.g. congenital adrenal hyperplasia); or extreme variations of internal and/or external genital organs (e.g. uterus didelphys or micropenis). In line with previous research, we also included more common variations of the external organs such as hypospadias, cryptorchism (for which an operation was needed), uterus anomalies or a vaginal septum as intersex conditions.[10] Although such common variations are generally not considered expressions of intersex conditions, we excluded all participants with any variations in this stereotype appearance, so that we include the most stereotypic women and men from a biological viewpoint (see ***Appendix A.1*** for more details).

Complementing approaches were applied to identify participants with an intersex condition (***Appendix A.2***). Firstly, the subsample of participants who's DNA had passed the earlier described quality control (n=13,395) was selected, since this selection excluded participants with copy number variations of the sex chromosomes, SRY-gene abnormalities, or participants with an officially changed transgender identity. Secondly, text fields asking about disorders, birth defects and operations were searched for expressions of potential intersex variations, intersex birth defects and sex-related operations, respectively. Lastly, we considered all biological males who used prescribed estrogens and all biological females who used prescribed testosterone transgender in the current study. Ultimately, 74 participants within the subsample were labeled as 'highly likely' of having an intersex condition or non-conform gender identity and were excluded (***Appendix A.3***). This resulted in a subsample of n=13,321 to construct a gender index.

Gender-related variables

All psychological and social variables are included in the model to construct the gender index, as far as these meet both of the following criteria: i) the variable is not reflecting a momentary emotional state that strongly fluctuates over short time periods and ii) the variable has <40% missing values. Since the included questionnaires were not originally constructed to identify potential psychosocial differences according to sex, potential sex-related differences are more likely to be found in single items, than in sumscores of questionnaires. Therefore, we included item-level variables. Only for the NEO-PI-R,[26] which assesses personality traits, we included mean subscale scores, rather than item scores, because we did not expect individual items to differ more between sexes than subscales would. This resulted in 153 variables, consisting of 145 single variables and 8 personality subscales, included in the analyses. These variables cover three of the four gender aspects which were previously defined, namely gender roles, gender dynamics and institutionalized gender.[9,27] ***Appendix B*** provides information on included variables.

Statistical analyses: Elastic net regularized generalized linear model

Based upon visual inspection of missing data patterns of the adult Lifelines population with the VIM Package 4.8.0, we concluded that there is no strong indication of data missing not at random and therefore multiple imputation with Mice Package 3.3.0 in R Studio 1.1.383 was performed.[28] Age, municipally-registered sex and source of entry into the study were always included as predictor variables for the multiple imputation. In total, 73% of all participants had at least one missing value on a variable that was relevant to construct the gender index. The variable with the highest frequency of missing data (20.9% of all participants) was membership of a social club. The minimal correlation of potential predictors with the variable to be imputed was 0.1. To construct a gender index, 245 variables (derived from 153 unique psychosocial variables potentially related to sex) were entered into an elastic net regularized generalized linear model with sex as dichotomous outcome (n=13,321).[29] This method retains parsimonious number of variables, which are highly predictive for the outcome and attains a high predictive accuracy of the model.[29,30] The data was randomly assigned to a training set (80%; n=10,657), and a testing set (20%, n=2,664). The former set was used to estimate coefficients: larger estimated coefficients indicate greater importance in discriminating between sexes. The latter set was used to calculate the model's predictive accuracy. The optimal regularization parameter α was selected by a grid search with the same 10-fold cross-validation for three αs, namely 0.1, 0.5 and 1.0. For the predictive model with the optimal α, the value of λ that minimized the mean squared error (MSE) was selected by 10-fold cross-validation, as was λ plus one standard error: $\lambda 1se$. The area under the receiving operating characteristic-curve (AUC) was calculated when predicting

classification of participants into 'male' or 'female' as the measure of goodness of fit. To provide an overview of the most discriminative gender-related variables, a model with the AUC of 0.80, generally already interpreted as 'good' classifying accuracy, was calculated as well in the test set.

Estimates of the coefficients obtained through the aforementioned regression ultimately formed the basis of the composite gender index that was applied to each adult Lifeline participant (n=152,728). The gender index is a continuum, ranging from 0% to 100%, representing the probability of each individual being a woman: the higher the gender index, the more feminine characteristics a person has. Androgyny is indicated by an index of 50%, where equal levels of feminine and masculine characteristics are present.

Statistical analyses: Analyses of common somatic symptoms and chronic diseases
Previous studies showed that age and educational levels are associated with gender roles,[31,32] thus we performed two-way ANOVAs to assess whether the magnitude of difference in gender index between sexes was equal across age groups (18-44, 45-64, 65+) and educational levels (low, medium, high) in the current study.[33,34]

Twelve multiple ordinal regressions were conducted to investigate whether sex and the gender index independently have an association with common somatic symptoms in the general population. Assumptions of ordinal regression were met.[35] Common somatic symptoms were assessed with the 12-item ordinal Symptom CheckList-90 Somatization subscale (SCL-90 SOM), which has been recommended for large scale studies.[36] The 12 items had five Likert-response options. We used multiple linear regression to investigate the associations of the gender index and sex with the SCL-90 SOM sumscore.

To test the associations of sex and the gender index with the lifetime prevalence of chronic diseases, we performed twelve multiple logistic regression analyses. Diseases that were identified by the Dutch Ministry of Health, Welfare and Sport as causing the greatest loss of healthy life years per person in the Netherlands were identified amongst Lifelines participants.[37-39] At baseline, self-reported lifetime prevalence hereof was measured. Low-prevalent (e.g. dementia) and sex-specific diseases (e.g. pregnancy diabetes) were excluded from the analyses.[40] Validity of the logistic regression's linearity assumption was violated for cardiovascular diseases (CVD; encompassing arrhythmia, heart failures and heart attack), skin cancer, epilepsy and asthma/COPD. Thus, gender indexes were categorized into quartiles that were included in the respective models. As a sensitivity analysis, categorization of gender indexes by means of a split at the median yielded comparable results.

We tested interaction terms between the gender index and sex for significance. We assessed multicollinearity of the variables by the variance inflation factor (VIF). No problems with multicollinearity were found, as VIF was <5 in all analyses. For all the above-mentioned analyses, we provided analysis codes in OSF (https://osf.io/z9aw4/) to increase transparency of the study.

Results

The gender index

The gridsearch to select the optimal regularization parameter had the best binomial deviance and minimal mean-squared errors (MSE) when α=1.0, equaling a LASSO regression.[30] For the predictive model with α=1.0, 10-fold cross validation selected the value of λ that minimized MSE (λ=9.7E-4; $\lambda.1se$ =2.4E-3). This was the sparsest model, with an accuracy comparable with the best predictive model. The model's AUC was 92% and the obtained coefficients were used as the basis for the gender index. Of the initial 245 potentially sex-related variables (representing 153 unique variables), 92 were excluded from the model and 153 (dummy) variables representing 85 unique variables remained. Many variables were highly indicative of sex. For reasons of clarity all estimated coefficients of nominal and ordinal variables with an OR below 0.5 or above 1.5 and all continuous variables are presented in ***Appendix C***. Most profound were physical activity-related variables (e.g. type of sport activities), work-related variables (e.g. profession), lifestyle (e.g. alcoholic uptake), tasks at home (e.g. cooking and household activities) and personality characteristics. In the model with an AUC of 80% (α=1.0; λ=0.12) nine variables remained, related to hours of work, hours of household activities including cooking dinner, and spending leisure time by performing odd jobs (***Table 2***).

The distribution of the gender index (range: 0-100% feminine) was bimodal. We found median gender indexes of 0.06 (IQR: 0.01, 0.23) for men, and 0.96 (IQR: 0.83, 0.99) for women. In 10.1% (n=15,480) of the participants the gender index was not in line with their biological sex: 12.5% (n=7,935) of all men scored 50% or higher on the gender index (indicating psychosocial femininity), and 8.4% (n=7,545) of all women scored less than 50% (indicating psychosocial masculinity). Significant interaction terms in two-way ANOVAs showed that the magnitude of the difference in gender index between men and women differed across age groups and educational level, thus analyses were adjusted for age and education.

Table 2. Predictors included in the model correctly classifying 80% of the participants' sex.

Predictor (ordered from strong to less strong)	Odds of being a woman[a]
Always preparing your own dinner	1.30
Days per week light to moderate household activities (0-7)	1.06
Hours per day light to moderate household activities (0-16)	1.06
Hours per week working (0-60+)	0.99
Days per week of leisure time spend on odd jobs (0-7)	0.99
Hours of leisure time spend on odd jobs (0-12)	0.96
Sometimes preparing your own dinner	0.80
Dinner is always prepared by someone else	0.65
Spending leisure time on odd jobs of light to moderate intensity	0.65

[a]Please note that the odds presented for the continuous predictor variables are per unit change on the scale of the predictor and are thus not directly comparable. ORs below 1.0 are indicative of being a male.

We performed our analysis with the gender index that was built upon a population that excluded people with an intersex condition. However, many large cohort studies do not include genetic data in addition to health measures. Thus, participants with intersex conditions in such cohorts cannot be excluded based on a genetic profile. Therefore, we also calculated a gender index based on the complete adult Lifelines population. We found median gender indexes of 0.05 (0.01, 0.22) for men, and 0.97 (0.87, 0.99) for women. A Wilcoxon Signed Rank test for related samples found no statistically significant difference for the gender indexes in both men and women Additionally, the two gender indexes are highly correlated (ρ>0.95).

Sex and gender, and the association with common somatic symptoms

Sex and gender (i.e. feminine or masculine characteristics) independently associated with common somatic symptoms (***Table 3***). Significant interaction terms in 9 of the 12 tested symptoms showed that the association of feminine gender characteristics with common somatic symptoms significantly differed per sex for the majority of symptoms. Therefore, we stratified the associations of feminine gender on the common somatic symptoms per sex.

Figure 1 shows that men's experiences of the common somatic symptoms are strongly associated with an increase in feminine gender characteristics. As suggested by all ORs exceeding 1.0, displaying feminine characteristics is associated with a higher common somatic symptom burden of all types. For example, a one hundred-percent increase in femininity score in men, is associated with 1.74 times higher odds on experiencing a one-point increase in severity of dizziness as measured by the SCL-90

SOM. In women, however, feminine characteristics are associated with less chest pain and less difficulties in breathing. The association between feminine characteristics and the sumscore of the SCL-90 SOM was significantly stronger in men than in women. In men, every one hundred-percent increase in feminine characteristics is associated with a 0.09-point increase in the SCL-90 SOM sumscore (β=0.09; 95% confidence interval: 0.08,0.10), whereas in women a one hundred-percent increase in feminine characteristics is associated to a 0.02-point increase in the SCL-90 SOM sumscore (β=0.02; 95% confidence interval: 0.01,0.03).

Table 3. The adjusted associations of sex and gender with common somatic symptoms.

Predictor	**Sex - female**		**Gender[a] - feminine**	
	OR 95% CI		**OR 95% CI**	
Outcome				
Headache	1.95	1.88, 2.03	1.34	1.28, 1.40
Dizziness	1.61	1.53, 1.69	1.33	1.25, 1.42
Chest pain	0.92	0.86, 0.99	*1.01*	*0.93, 1.10*
Lower back pain	1.24	1.20, 1.29	1.09	1.04, 1.14
Nausea	1.49	1.42, 1.56	1.24	1.17, 1.32
Painful muscles	1.29	1.24, 1.34	1.07	1.03, 1.12
Difficulties breathing	1.16	1.09, 1.24	*1.02*	*0.94, 1.10*
Feeling hot/cold	3.21	3.05, 3.37	1.24	1.17, 1.31
Numbness or tingling	1.14	1.09, 1.20	1.16	1.09, 1.22
Feeling lump in throat	1.62	1.53, 1.72	1.28	1.20, 1.38
Weakness body parts	*1.03*	*0.99, 1.08*	1.38	1.31, 1.46
Heavy arms or legs	1.27	1.21, 1.33	1.21	1.15, 1.28
	B coefficient 95% CI		**B coefficient 95% CI**	
SCL90SOM sumscore	0.09	0.08, 0.10	0.05	0.04, 0.06

[a]Gender as indicated by the calculated gender index, ranging from 0% (masculine) to 100% (feminine). NB: The association of sex with common somatic symptoms is adjusted for gender, age and education, whereas the association of gender with common somatic symptoms is adjusted for sex, age and education.

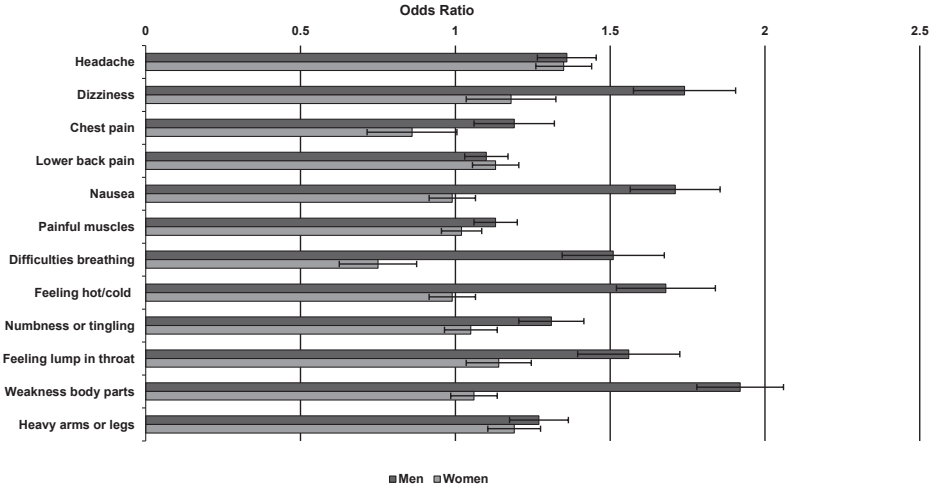


Figure 1: The associations of feminine gender with reported common somatic symptoms, stratified by sex.

Sex and gender, and the association with chronic diseases

Table 4 shows that sex and the gender index have an independent association with the lifetime prevalence of most chronic diseases. An increase in gender index, i.e. displaying more feminine characteristics, is associated with an increased risk of chronic diseases. An example hereof is the association between a hundred-percent increase in feminine characteristics and the 1.51-times higher odds on having a stroke. In contrast, female sex appears to be protective of most chronic diseases, except for osteoarthritis, migraine and osteoporosis. Significant interaction terms in five of 12 tested chronic diseases showed that the associations of femininity with lifetime prevalence of chronic diseases significantly differed per sex. Therefore, we stratified the association of femininity with lifetime prevalence of chronic diseases per sex.

Figure 2 shows that in women increases in gender index appeared to be not significantly associated with the lifetime prevalence of chronic diseases, except for migraine (OR=1.25; 95% confidence interval: 1.16, 1.36). In men, however, increases in feminine characteristics were found to be associated with an increased lifetime prevalence of most chronic diseases.

Sensitivity analyses

To explore the possibility that the associations were identified due to respondents' health status, rather than gender roles, a sensitivity analysis was performed by excluding physical activity-related predictors. The analysis showed that composition

of the gender index remained similar, and was highly correlated to the original gender index (ρ>0.95) as well. In addition, the associations of the gender index without physical activity-related variables with common somatic symptoms and chronic diseases remained similar to those found with the original gender index. To explore the influence of the imputations, we performed a second sensitivity analysis based on respondents with complete data. The analyses showed that again that the composition of the gender index remained mostly comparable and was highly correlated to the original gender index (ρ=0.95). Additionally, the associations of the gender index with common somatic symptoms and chronic diseases remained comparable. Results of both sensitivity analyses are provided in the electronic supplementary materials.

Table 4. The adjusted associations of sex and gender with the prevalence of chronic disease.

Predictor	**Sex – female**		**Gender[a] – feminine**	
	OR	**95% CI**	**OR**	**95% CI**
Outcome				
CVD	0.88	0.82, 0.91	1.17	1.09, 1.26
Epilepsy	0.66	0.57, 0.78	1.60	1.34, 1.91
Asthma/COPD	*1.01*	*0.95, 1.06*	1.08	1.02, 1.15
DM1	0.54	0.37, 0.79	*1.44*	*0.91, 2.27*
DM2	0.48	0.42, 0.55	1.91	1.61, 2.26
Stroke	0.57	0.46, 0.71	1.51	1.16, 1.97
Osteoarthritis	1.68	1.55, 1.82	*1.09*	*0.98, 1.20*
Skin cancer	*1.16*	*0.99, 1.35*	*1.13*	*0.96, 1.34*
Kidney failure	*1.08*	*0.95, 1.22*	1.19	1.02, 1.38
Migraine	2.37	2.25, 2.49	1.27	1.19, 1.35
Osteoporosis	4.73	3.89, 5.75	*1.07*	*0.86, 1.33*
Rheumatoid arthritis	*1.04*	*0.91, 1.19*	1.55	1.32, 1.83

[a]Gender as indicated by the calculated gender index, ranging from 0% (masculine) to 100% (feminine). NB: The linearity assumption for logistic regression between gender and CVD, skin cancer, epilepsy and asthma/COPD was violated. Therefore, results here are the odds of the highest gender quartile compared to the lowest gender quartile. The associations of sex with the prevalence of chronic diseases is adjusted for gender, age and education, whereas the association of gender with the prevalence of chronic disease is adjusted for sex, age and education.

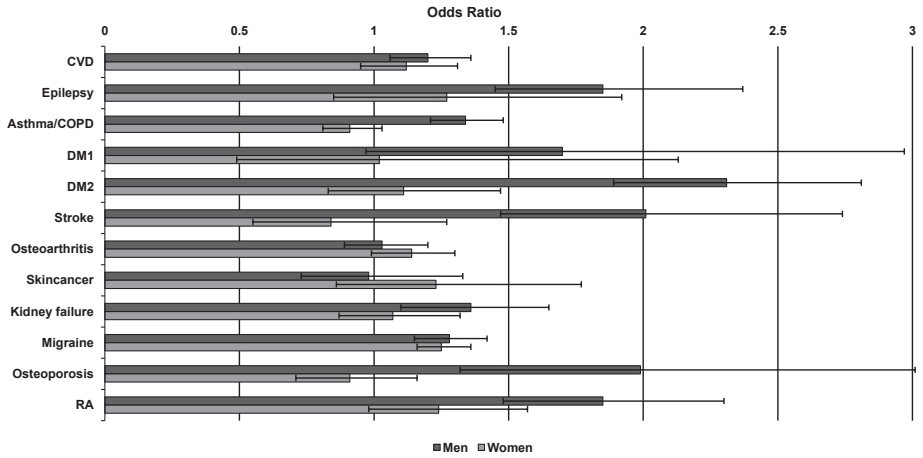


Figure 2: The associations of feminine gender with reporting chronic diseases, stratified by sex.

Discussion

We found that a higher gender index (i.e. tending towards feminine characteristics) and sex are independently associated with common somatic symptoms and lifetime prevalence of chronic diseases. Feminine characteristics are associated with experiencing a higher common somatic symptom burden and chronic diseases, especially in men. Female sex is also associated with a higher burden of common somatic symptoms, but not to all chronic diseases, compared to male sex. In 10.1% of the participants, the gender index was not in line with participants' sex, as in 12.5% of men and 8.4% of women a discrepancy between gender index and sex was found.

Strengths and limitations

This study has some strengths. First, LifeLines includes a large sample size in which many psychosocial variables were assessed. This allowed us to incorporate a wide range of gender-related psychosocial candidate variables into the model. Furthermore, it assured strong statistical power. Second, the current method of calculating a gender index follows a data-driven approach, whereas most previously established gender indexes follow a theory-driven approach.[8,19,27,41] Both approaches have their respective advantages and disadvantages. A data-driven approach allows for a model to include a wide range of psychosocial characteristics to construct a gender index and to move beyond merely operationalizing gender roles by means of personality characteristics. Furthermore, it allows to establish a time- and place-dependent gender index since the index is constructed based on psychosocial differences between biological sexes

in the cohort under study. In other words, the method is flexible and adaptive and can be applied to calculate a gender index specific to the study's context. Other studies do not necessarily need to rely on genetic data of participants to exclude people with indeterminate binary sex, as we showed that excluding the people with an intersex condition from a large general population cohort did not change the constructed gender index significantly. However, one should note that the estimated coefficients are chosen merely based on maximizing predictive accuracy of a model and the nature of the predictors is ignored. That is, when two variables are equally contributing to improving the model's predictive accuracy, the algorithm randomly chooses one of the predictors. Additionally, the gender index proposed here is a study-specific measure for gender roles that cannot be generalized to other studies: although the method is applicable in other studies, the exact construct of the gender index will vary across settings.

A theory-driven gender index allows for easier comparison across studies and its interpretation is more straightforward than that of a data-driven gender index. However, it should be noted that a theory-driven gender index cannot handle the time- and place-dependent nature of (the embodiment of) gender roles, i.e. the relevance of the index might differ between studies performed at different times or at different places. Given these advantages and disadvantages of both approaches, a combination of a theory-driven approach (in which potential gender-related variables are selected based on theory or expert opinion) and a data-driven approach (in which an algorithm maximizes the predictive accuracy of the model underlying the gender index) may yield the easiest interpretable and most trustworthy gender index.

Limitations of the study should be acknowledged as well. First, we did not include a gold standard gender measure, as this does not exist yet.[42] Thus, our gender index was not validated with other measures. Second, this study is not gender-expansive, as it does not move beyond the male/female binary. Therefore, our results are not directly applicable to agender or non-binary people. Third, lifetime prevalence of chronic diseases was self-reported, which could cause response and recall bias. These biases might differ between women and men. Several studies suggest that men are less willing to report health problems and report health problems at a later time than women.[43,44] Therefore, the reported association of female sex -and by inference that of femininity- with chronic diseases might be an overestimation. Last, this study had a cross-sectional design, which does not allow for any causal inference of the effect of gender on common somatic symptoms or chronic diseases, or vice versa. Therefore, no causality can be concluded from the study. In addition, we cannot exclude the possibility that the reported associations are partly spurious, since the

predictors on which the gender index is based may not only reflect respondents' gender roles but also their general health. However, a sensitivity analyses using a gender index that was estimated without physical activity-related variables showed that the associations of sex and gender with health outcomes remained comparable, suggesting that the reported associations are largely meaningful.

Sex and gender, and the relation to common somatic symptoms and chronic disease
To the authors' knowledge, this is the first study in which gender indexes were derived in a data-driven manner, i.e. without pre-specified variables. The defined model displayed 92% accuracy in distinguishing between sexes, indicating that the included psychosocial variables are strongly connected to being a man or a woman in the LifeLines Cohort. The variables with the highest predictive value for sex (i.e. work-related and household-related variables, and dedicating leisure time to odd jobs) in the current study, were in concordance with previous studies.[8,19,27] A newly identified variable within the realm of household-related variables that strongly discriminates between sexes included the frequency in which one prepares his or her own dinner.

We found that household-related activities had strong predictive value in discriminating between sexes. A Swedish study found 2.2-fold higher odds on experiencing common somatic symptoms for women who had many domestic duties, compared to women with few domestic duties.[45] Follow-up studies replicated this: women with jobstrain and domestic duties had higher odds on experiencing common somatic symptoms than women with jobstrain and no domestic duties.[46,47] These studies support the idea that female gender roles are more stressful and less gratifying than male gender roles, which might lead to worse health outcomes.[48] Household responsibilities have been associated with lower access to health care in both men and women,[49] suggesting that healthcare factors might explain health-related gender differences as well. Finally, we cannot exclude reverse causality, in which experiencing common somatic symptoms results in less paid working hours and therefore conducting more household-related activities.[50]

Literature on the association of gender with common somatic symptoms and chronic diseases is scarce. Although many studies do not explicitly distinguish between sex and gender in health,[51,52] some studies suggest that differences in gender roles are mediating sex differences in experienced health.[53-55] The current study also shows that female sex and feminine gender roles are independently associated with increased common somatic symptoms. Earlier research has shown that female sex is a risk factor for a variety of symptoms.[6,15] Our findings show that female sex compared to male sex is associated with an increased lifetime prevalence of osteoporosis, migraine

and osteoarthritis, which is in line with previous studies.[12,56,57] The current results also suggest that male sex was associated with an increased prevalence of CVD, which is in line with the traditional, yet increasingly defeated, idea of CVD being '*a man's disease*'.[58,59] However, it must be noted that underdiagnoses of -amongst others- cardiovascular diseases, still tends to occur in women, despite improvements made over the last decade.[58] Additionally, women might undergo different (diagnostic) care trajectories than men, rendering women with more delayed or wrong diagnoses.[60,61] This underreporting might partly explain the protective association of female sex with cardiovascular disease.

To the authors' knowledge, this is the first study to show that feminine gender is associated with an increased burden of common somatic symptoms and a higher lifetime prevalence of chronic diseases, especially in men. This is in line with previous studies that showed that higher frequencies of common somatic symptoms are reported in gender dysphoric (i.e. people in whom sex assigned at birth and current gender identity do not match) and gender non-conforming individuals, than in cisgender individuals.[62-65] More specifically, significantly higher SCL-90 SOM scores were reported in male-to-female transsexual and transgender individuals than in cisgender men.[66] These increased rates of common somatic symptoms are often attributed to the psychological distress that is inherent to gender dysphoria. The inability to adhere to imposed societal norms on masculinity and femininity is theorized to cause anticipated and internalized stigma in these individuals, which results in a higher chance of reporting somatic symptoms.[64] Although the feminine men in our study are not necessarily gender dysphoric (they merely display more feminine than masculine characteristics) the psychological distress that accompanies the non-adherence to societal norms on gender roles together with the stress that accompanies female gender roles could have affected these participants as well. Only one previous study amongst the general population tested whether femininity was associated with common somatic symptoms. This study found no significant association, but had a relatively small sample size and used the BSRI.[67] No distinction between sexes was made in this study.

In conclusion, incorporating one's gender roles and sex in care trajectories could aid the process of effective medicine, tailored to the societal circumstances in which gender roles are shaped. Therefore, we recommend to conduct further research to explore the effect of gender on health outcomes in clinical settings. In line with this, research could focus on the association of gender with established risk factors for disease, instead of the disease itself, for this might allow for more effective preventative interventions. Additionally, a longitudinal study design could be useful

to explore possible reverse causality between gender roles and health and to obtain insight into the relation between the two concepts. We also suggest to set a stricter distinction between gender and biological sex in research and literature, as these concepts are often applied interchangeably. Lastly, we recommend the consideration of gender in large cohort studies, as our methodology shows this is feasible.

References

1. Pardue M, Wizemann TM. *Exploring the biological contributions to human health: Does sex matter?* Washington DC: National Academies Press; 2001.
2. Phillips SP. Including gender in public health research. *Public Health Rep*. 2011;126(3):16-21. doi: 10.1177%2F00333549111260S304.
3. Springer KW, Stellman JM, Jordan-Young RM. Beyond a catalogue of differences: A theoretical frame and good practice guidelines for researching sex/gender in human health. *Social Science & Medicine*. 2012;74(11):1817-1824. doi: 10.1016/j.socscimed.2011.05.033.
4. Labaka A, Goñi-Balentziaga O, Lebeña A, Pérez-Tejada J. Biological sex differences in depression: A systematic review. *Biological research for nursing*. 2018;20(4):383-392. doi: 10.1177/1099800418776082.
5. Spence JD, Pilote L. Importance of sex and gender in atherosclerosis and cardiovascular disease. *Atherosclerosis*. 2015;241(1):208-210. doi: 10.1016/j.atherosclerosis.2015.04.806 [doi].
6. Tomenson B, Essau C, Jacobi F, et al. Total somatic symptom score as a predictor of health outcome in somatic symptom disorders. *The British Journal of Psychiatry*. 2013;203(5):373-380. doi: 10.1192/bjp.bp.112.114405.
7. Nowatzki N, Grant KR. Sex is not enough: The need for gender-based analysis in health research. *Health Care for Women International*. 2011;32(4):263-277. doi: 10.1080/07399332.2010.519838.
8. Smith PM, Koehoorn M. Measuring gender when you don't have a gender measure: Constructing a gender index using survey data. *International Journal for Equity in Health*. 2016;74(6):459-468. doi: 10.1186/s12939-016-0370-4.
9. Johnson JL, Greaves L, Repta R. *Better science with sex and gender: A primer for health research*. Vancouver: Women's Health Research Network Vancouver; 2007.
10. Blackless M, Charuvastra A, Derryck A, Fausto-Sterling A, Lauzanne K, Lee E. How sexually dimorphic are we? review and synthesis. *American Journal of Human Biology*. 2000;12(2):151-166.
11. Bottorff JL, Oliffe JL, Robinson CA, Carey J. Gender relations and health research: A review of current practices. *International Journal for Equity in Health*. 2011;10:60. doi: 10.1186/1475-9276-10-60.
12. Alswat KA. Gender disparities in osteoporosis. *Journal of Clinical Medicine Research*. 2017;9(5):382-387. doi: 10.14740/jocmr2970w.
13. Noh JW, Park H, Kim M, Kwon YD. Gender differences and socioeconomic factors related to osteoporosis: A cross-sectional analysis of nationally representative data. *Journal of Women's Health*. 2018;27(2). doi: 10.1089/jwh.2016.6244.
14. Galdas PM, Cheater F, Marshall P. Men and health help-seeking behaviour: Literature review. *J Adv Nurs*. 2005;49(6):616-623. doi: 10.1111/j.1365-2648.2004.03331.x.
15. Barsky AJ, Peekna HM, Borus JF. Somatic symptom reporting in women and men. *Journal of General Internal Medicine*. 2001;16(4):266-275. doi: 10.1046/j.1525-1497.2001.00229.x.
16. Lee C, Owens RG. Issues for a psychology of men's health. *Journal of Health Psychology*. 2002;7:209-17.
17. Tannenbaum C, Ellis RP, Eyssel F, Zou J, Schiebinger L. Sex and gender analysis improves science and engineering. *Nature*. 2019;575(7781):137-146. doi: 10.1038/s41586-019-1657-6.
18. Helmreich RL, Spence JT, Wilhelm JA. A psychometric analysis of the personal attributes questionnaire. *Sex Roles*. 1981;7(11):1097-1108.

19. Lippa R, Connelly S. Gender diagnosticity: A new bayesian approach to gender-related individual differences. *Journal of Personality and Social Psychology*. 1990;59(5):1051-1065. doi: 10.1037/0022-3514.59.5.1051.
20. Hoffman RM, Borders LD. Twenty-five years after the bem sex-role inventory: A reassessment and new issues regarding classification variability. *Measurement and Evaluation in Counseling and Development*. 2001;34(1):39-55. doi: 10.1080/07481756.2001.12069021.
21. Choi N, Fuqua DR, Newman JL. The bem sex-role inventory: Continuing theoretical problems. *Educational and Psychological Measurement*. 2008;68(5):881-900. doi: 10.1177/0013164408315267.
22. Pedhazur EJ, Tetenbaum TJ. Bem sex role inventory: A theoretical and methodological critique. *Journal of Personality and Social Psychology*. 1979;37(6):996-1016. doi: 10.1037/0022-3514.37.6.996.
23. Klijs B, Scholtens S, Mandemakers JJ, Snieder H, Stolk RP, Smidt N. Representativeness of the LifeLines cohort study. *PLoS ONE*. 2015;10(9):e0137203. doi: 10.1371/journal.pone.0137203.
24. Scholtens S, Smidt N, Swertz MA, et al. Cohort profile: LifeLines, a three-generation cohort study and biobank. *International Journal of Epidemiology*. 2015;44(4):1172-80. doi: 10.1093/ije/dyu229.
25. Furtado PS, Moraes F, Lago R, Barros LO, Toralles MB, Barroso U. Gender dysphoria associated with disorders of sex development. *Nature Reviews Urology*. 2012;9(11):620-627. doi: 10.1038/nrurol.2012.182.
26. Costa PT, McCrae RR. Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI) professional manual. *Psychological Assessment Resources*. 1992.
27. Pelletier R, Ditto BF, Pilote L. A composite measure of gender and its association with risk factors in patients with premature acute coronary syndrome. *Psychosomatic medicine*. 2015;77(5):517-526. doi: 10.1097/PSY.0000000000000186.
28. van Buuren, S., Groothuis-Oudshoorn,K. Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*. 2011;45(3). doi: 10.18637/jss.v045.i03.
29. Zou H, Hastie T. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society.* 2005;67(2):301-320.
30. Hastie T, Qian J. Glmnet vignette. https://web.stanford.edu/~hastie/glmnet/glmnet_alpha.html#log. Updated 2014. Accessed April 25th, 2019.
31. Koenig AM. Comparing prescriptive and descriptive gender stereotypes about children, adults, and the elderly. *Frontiers in Psychology*. 2018;9:1086. doi: 10.3389/fpsyg.2018.01086.
32. Miller J, Chamberlin M. Women are teachers, men are professors: A study of student perceptions. *Teaching Sociology*. 2000;28(4):283-298. doi: 10.2307/1318580.
33. van Zon SKR, Snieder H, Bültmann U, Reijneveld SA. The interaction of socioeconomic position and type 2 diabetes mellitus family history: A cross-sectional analysis of the lifelines cohort and biobank study. *BMJ Open*. 2017;7(4):e015275. doi: 10.1136/bmjopen-2016-015275.
34. Hilderink PH, Collard R, Rosmalen JGM, Oude Voshaar RC. Prevalence of somatoform disorders and medically unexplained symptoms in old age populations in comparison with younger age groups: A systematic review. *Ageing Research Reviews*. 2013;12(1):151-156.
35. O'Connell AA. *The cumulative (proportional) odds model for ordinal outcomes.* Thousand Oaks, CA: SAGE Publications, Inc.; 2006. 10.4135/9781412984812.
36. Zijlema WL, Stolk RP, Löwe B, Rief W, White PD, Rosmalen JGM. How to assess common somatic symptoms in large-scale studies: A systematic review of questionnaires. *Journal of Psychosomatic Research*. 2013;74(6):459-68. doi: 10.1016/j.jpsychores.2013.03.093.

37. Dutch Ministry of Health, Welfare and Sports. Ranglijst aandoeningen op basis van verlies aan gezonde levensjaren (ziektejaarequivalenten). https://www.volksgezondheidenzorg.info/ranglijst/ranglijst-aandoeningen-op-basis-van-verlies-aan-gezonde-levensjaren-ziektejaarequivalenten. Updated 2019. Accessed May 23rd, 2019.
38. Hoeymans N, Gijsen R, Slobbe LC. 59 important health problems; a selection of diseases for public health monitoring. *Nederlands tijdschrift voor geneeskunde*. 2013;157(31):A5994.
39. Gijsen R, Poos MJJC, Slobbe LC, Mulder M, In 't Panhuis-Plasmans MHD, Hoeymans N. Een nieuwe selectie van ziekten voor de volksgezondheid toekomst verkenningen. *RIVM*. 2013;010003004/2013.
40. Green SB. How many subjects does it take to do A regression analysis. *Multivariate Behavioral Research*. 1991;26(3):499-510.
41. Bem SL. The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*. 1974;42(2):155-162. doi: 10.1037/h0036215.
42. Schiebinger L, Stefanick ML. Gender matters in biological research and medical practice. *Journal of the American College of Cardiology*. 2016;67(2):136-138. doi: 10.1016/j.jacc.2015.11.029.
43. Singh-Manoux A, Guéguen A, Ferrie J, et al. Gender differences in the association between morbidity and mortality among middle-aged men and women. *American Journal of Public Health*. 2008;98(12):2251-2257. doi: 10.2105/AJPH.2006.107912.
44. Hausteiner-Wiehle C, Schneider G, Lee S, Sumathipala A, Creed F. Gender, lifespan and cultural aspects. In: Creed F, Henningsen P, Fink P, eds. *Medically unexplained symptoms, somatisation and bodily distress.* Cambridge, UK: Cambride University Press; 2011:132-155.
45. Krantz G, Ostergren P. Double exposure: The combined impact of domestic responsibilities and job strain on common symptoms in employed swedish women. *The European Journal of Public Health*. 2001;11(4):413-419.
46. Mellner C, Krantz G, Lundberg U. Symptom reporting and self-rated health among women in mid-life: The role of work characteristics and family responsibilities. *International Journal of Behavioral Medicine*. 2006;13(1):1-7.
47. Krantz G, Berntsson L, Lundberg U. Total workload, work stress and perceived symptoms in swedish male and female white-collar employees. *The European Journal of Public Health*. 2005;15(2):209-214.
48. Nathanson CA. Illness and the feminine role: A theoretical review. *Social Science & Medicine*. 1975;9(2):57-62. doi: 10.1016/0037-7856(75)90094-3.
49. Pelletier R, Humphries KH, Shimony A, et al. Sex-related differences in access to care among patients with premature acute coronary syndrome. *Canadian Medical Association Journal*. 2014;186(7):497-504. doi: 10.1503/cmaj.131450.
50. van der Leeuw G, Gerrits MJ, Terluin B, et al. The association between somatization and disability in primary care patients. *Journal of Psychosomatic Research*. 2015;79(2):117-122. doi: 10.1016/j.jpsychores.2015.03.001.
51. Regitz-Zagrosek V. Sex and gender differences in health. *EMBO reports*. 2012;13(7):596-603. doi: 10.1038/embor.2012.87.
52. Hankivsky O, Springer KW, Hunting G. Beyond sex and gender difference in funding and reporting of health research. *Research integrity and peer review*. 2018;3(1):6.
53. Robinson ME, Riley JL, Myers CD, et al. Gender role expectations of pain: Relationship to sex differences in pain. *The Journal Of Pain*. 2001;2(5):251-257.

54. Alabas OA, Tashani OA, Tabasam G, Johnson MI. Gender role affects experimental pain responses: A systematic review with meta-analysis. *European Journal of Pain*. 2012;16(9):1211-1223. doi: 10.1002/j.1532-2149.2012.00121.x.
55. Pelletier R, Khan NA, Cox J, et al. Sex versus gender-related characteristics: Which predicts outcome after acute coronary syndrome in the young? *Journal of the American College of Cardiology*. 2016;67(2):127-135. doi: 10.1016/j.jacc.2015.10.067.
56. Hame SL, Alexander RA. Knee osteoarthritis in women. *Current reviews in musculoskeletal medicine*. 2013;6(2):182-187. doi: 10.1007/s12178-013-9164-0.
57. Schroeder RA, Brandes J, Buse DC, et al. Sex and gender differences in Migraine—Evaluating knowledge gaps. *Journal of Women's Health*. 2018;27(8):965-973. doi: 10.1089/jwh.2018.7274.
58. Wenger NK. Women and coronary heart disease: A century after herrick: Understudied, underdiagnosed, and undertreated. *Circulation*. 2012;126(5):604-611. doi: 10.1161/circulationaha.111.086892.
59. Humphries KH, Izadnegahdar M, Sedlak T, et al. Sex differences in cardiovascular disease - impact on care and outcomes. *Front Neuroendocrinol*. 2017;46:46-70. doi: 10.1016/j.yfrne.2017.04.001.
60. Arber S, McKinlay J, Adams A, Marceau L, Link C, O'Donnell A. Influence of patient characteristics on doctors' questioning and lifestyle advice for coronary heart disease: A UK/US video experiment. *The British Journal of General Practice*. 2004;54(506):673-678.
61. Arber S, McKinlay J, Adams A, Marceau L, Link C, O'Donnell A. Patient characteristics and inequalities in doctors' diagnostic and management strategies relating to CHD: A video-simulation experiment. *Social Science & Medicine*. 2006;62(1):103-115.
62. Heylens G, Verroken C, De Cock S, T'Sjoen G, De Cuypere G. Effects of different steps in gender reassignment therapy on psychopathology: A prospective study of persons with a gender identity disorder. *Journal of sexual medicine*. 2014;11(1):119-126. 10.1111/jsm.12363. doi: 10.1111/jsm.12363.
63. Colizzi M, Costa R, Todarello O. Dissociative symptoms in individuals with gender dysphoria: Is the elevated prevalence real? *Psychiatry Research*. 2015;226(1):173-180.
64. Serpe C. Somatization disorder and gender. In: Nadal KL, ed. *The SAGE encyclopedia of psychology and gender*. SAGE Publications; 2017:1606-1610.
65. Shirdel-Havar E, Steensma TD, Cohen-Kettenis PT, Kreukels BP. Psychological symptoms and body image in individuals with gender dysphoria: A comparison between iranian and dutch clinics. *International Journal of Transgenderism*. 2019;20(1):108-117.
66. Auer MK, Höhne N, Bazarra-Castro MA, et al. Psychopathological profiles in transsexuals and the challenge of their special status among the sexes. *PLoS One*. 2013;8(10):e78469. doi: 10.1371/journal.pone.0078469.
67. Castro Y, Carbonell JL, Anestis JC. The influence of gender role on the prediction of antisocial behaviour and somatization. *Int J Soc Psychiatry*. 2012;58(4):409-416. doi: 10.1177/0020764011406807.

Appendices

Appendix A.1. Potential Intersex Conditions

Copy number variations of the sex chromosomes

45, X0	46, XX/XY	47, XXY	47, XXX
48, XXXX	48, XXYY	49, XXXXX	49, XXXXY

Genetic mutations

5α-reductase deficiency	Congenital adrenal hyperplasia
17β-Hydroxysteroid dehydrogenase deficiency	Androgen insensitivity syndrome
Aromatase deficiency	Aromatase excess syndrome
Isolated 17,20-lyase deficiency	Leydig cell hypoplasia
Lipoid congenital adrenal hyperplasia	Swyer Syndrome/Gonadal dysgenesis

Developmental variations of internal and/or external sex organs

Aphallia	Diphallia
Micropenis	Cryptorchism
Müllerian agenesis (MRKH syndrome)	Clitoromegaly
Pseudovaginal perineoscrotal hypospadias	Uterus didelphys
Uterus unicornis	Additional uterine anomalies
Aposthia	Hypospadias

Appendix A.2. Sampling method

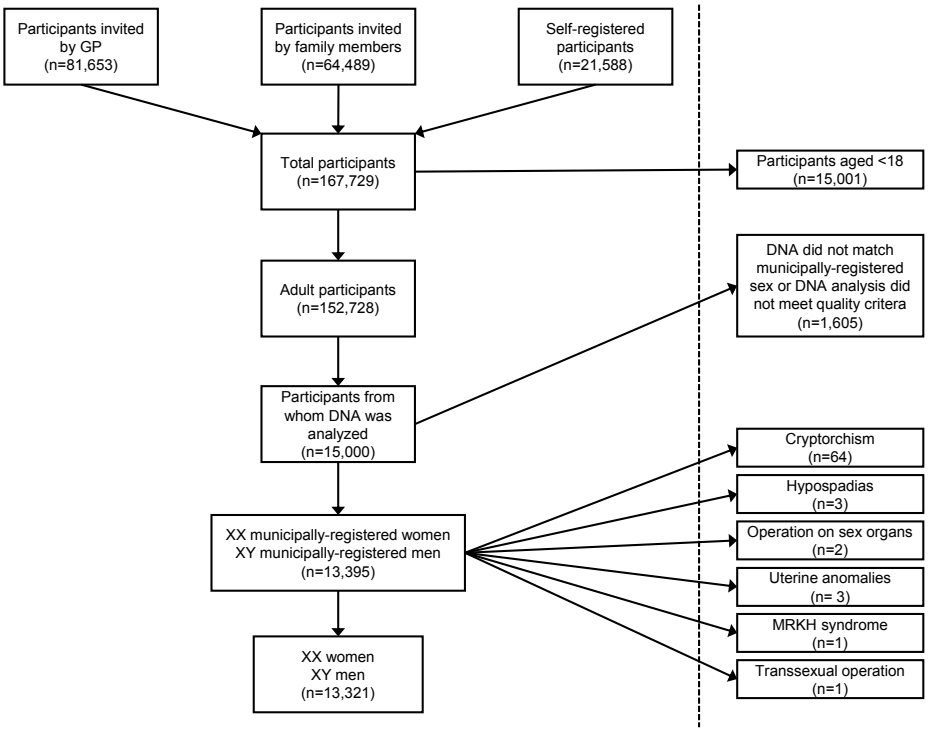


Figure A.2: Sampling method to identify participants with an intersex condition

Appendix A.3. Number of participants with potential intersex and/or non-conform gender identity, within the complete Lifelines sample and the subsample from whom DNA passed quality control.

Type of potential intersex condition and/or indication of transgender identity	Adult Lifelines Cohort (n=152,728)	Subsample (n=13,395)
Copy number variations of sex chromosomes		
Turner	11	0
Klinefelter	13	0
Genetic mutations (hormonal/metabolic)		
Congenital adrenal hyperplasia	8	0
Congenital variations of internal and/or external sex organs		
Hypospadia	53	3
Cryptorchism[a]	1,106	64
Operated for dysplastic or divergent sex organs[b]	42	2
Primary hypogonadism	10	0
Ovarian tube(s) missing or underdeveloped	7	0
Uterine anomalies	28	3
Müllerian agenesis (MRKH syndrome)	11	1
Gonadal dysgenesis (Swyer syndrome)	1	0
Non-conform gender identity		
Expressed gender dysphoria or non-conform gender identity and/or transgender medication use and/or transsexual operation(s)	19	1[c]
Total	1,309 (0.86%)	74 (0.55%)

[a]Cryptorchism included if present at adult age and/or reported as operated:99% of the 1106 participants who reported cryptorchism, reported it as operated. [b]e.g. ''no vagina'', ''vaginal septum'', ''divergent sex organs adjusted''. [c]The participant informed the researchers of the change in sex in the municipal registration and was thus not excluded during the quality control procedures of the genetic material.

Appendix B. Categories of variables included in LASSO regression

	Categories
1	General information and demographics
2	Current and past relationships
3	Living situation
4	Education and work
5	Social activities and wellbeing
6	Lifestyle
7	Diet and weight beliefs
8	Threatening experiences and long-term difficulties
9	Personality (NEO-PI-R)

Note: Detailed information on the included variables are included on OSF.

Appendix C. Nominal and ordinal predictors (odds >1.5 or <0.5) and all continuous predictors in the model with 92% predictive ability of the participants' sex.

Nominal and ordinal predictors (ordered from strong to less strong)	Odds of being a woman
Performing sport activity: horseriding	6.89
Performing sport activity: zumba[a]	5.58
Performing sport activity: soccer	0.19
Profession: crafts and related trades workers	0.21
Always preparing your own dinner	4.34
Being retired	0.26
Drinking alcohol 6-7 days a week	0.27
Doing leisure time odd jobs of moderate intensity	0.28
Dinner is always prepared by someone else	0.32
Drinking alcohol 3 days a week	0.32
Profession: plant and machine operators and assemblers	0.32
Currently in a job	2.99
Being a housewife or househusband	2.82
Short period of time dieting	2.56
Drinking alcohol 4-5 days a week	0.40
Doing leisure time odd jobs of light intensity	0.41
Drinking alcohol 1 day a week	0.42
Profession: skilled agricultural, forestry or fishery workers	0.44
Performing sport activity: gymnastics	2.26
Often preparing your own dinner	2.15
Losing one's job and not able to find new work	1.99

Appendix C. Continued.

Nominal and ordinal predictors (ordered from strong to less strong)	Odds of being a woman
Profession: services and sales workers	1.77
Unpleasant experience: got in trouble with the law or police in the past year	1.73
Profession: clerical support workers	1.55
Performing sport activity: swimming	1.54
Experienced difficulties and stress in the relationship with ones parents	1.53
Continuous predictors (odds per unit change on respective scale)	
Higher mean scores on the self-discipline scale of the NEO (range: 0-4)[b]	1.52
Higher mean scores on the impulsiveness scale of the NEO (range: 0-4)[b]	1.38
Higher mean scores on the self-consciousness scale of the NEO (range: 0-4)[b]	1.38
Number of household members smoking (range: 0-6+)	1.29
Hours per day light to moderate household activities (range: 0-16)	1.29
Higher mean scores on the vulnerability scale of the NEO (range: 0-4)[b]	1.29
Hours per day vigorous household activities (range: 0-16)	1.25
Hours sleep per 24 hours (range: 4-20)	1.19
Days per week light to moderate household activities (range: 0-7)	1.13
Higher scores on the competence scale of the NEO (range: 0-4)[b]	1.08
Days per week light to moderate household activities (range: 0-7)	1.04
Number of times moved house (range: 0-25+)	1.01
Percent declared unfit for work (range: 0-100%)	0.99
Days per week walking (range: 0-7)	0.98
Days per week at least 30 minutes light to moderate work (range: 0-7)	0.97
Number of cigars smoked per day (range: 0-10+)	0.97
Hours per day light to moderate work (range: 0-16)	0.96
Hours per week volunteering (range: 0-60)	0.95
Hours per week working (range: 0-60)	0.94
Number of co-residents (range: 0-6+)	0.93
Hours per day TV-watching (range: 0-8)	0.92
Hours per day cycling (0-12)	0.90
Higher mean scores on the deliberation scale of the NEO (range: 0-4)[b]	0.88
Hours per day odd jobs (range: 0-12)	0.86
Days of the week odd jobs (range: 0-7)	0.85
Higher mean scores on the hostility scale of the NEO (range: 0-4)[b]	0.80
Higher mean scores on the excitement scale of the NEO (range: 0-4)[b]	0.51

[a] Zumba is dance-based type of fitness. [b]Mean subscale scores of the NEO-PI-R. NB: An OR below 1.0 indicates being a man.

CHAPTER 5

Ballering, A.V., Wardenaar, K.J., olde Hartman, T.C., & Rosmalen, J.G.M. (2022). Female sex and femininity independently associate with common somatic symptom trajectories. *Psychological Medicine*, *52*(11), 2144-2154.

Abstract

Background: Multiple predictors have been associated with persistent somatic symptoms. However, previous studies problematically defined the persistence of symptoms, conflated participants' sex and gender, and focused on patient populations. Therefore, we studied associations between predictors, especially sex and gender, and longitudinal patterns of somatic symptoms in the general adult population. We also assessed whether predictors for persisting symptoms differ between sexes.

Method: To identify developmental trajectories of somatic symptoms, assessed by the SCL-90 SOM, we used latent class trajectory modeling in the Dutch Lifelines Cohort Study (N=150,494; 58.6% female; median time to follow-up: 46.0 [min-max: 22.0-123.0] months). To identify predictors of trajectories, we applied multiple logistic regression analyses. Predictors were measured by surveys at baseline and a composite gender index was previously developed.

Results: A five-class linear LCGA model fitted the data best: 93.7% of the population had a stable symptom trajectory, whereas 1.5% and 4.8% of the population had a consistently increasing or decreasing symptom trajectory, respectively. Female sex predicted severe, stable symptom severity (OR=1.74; 95%CI=1.36-2.22), but not increasing symptom severity (OR=1.15; 95%CI=0.99-1.40). Femininity was protective hereof (OR=0.60; 95%CI=0.44-0.82 and OR=0.66; 95%CI=0.51-0.85, respectively). Merely few predictors of symptom severity, for instance hours of paid employment and physical functioning, differed in strength between sexes. Yet, effect sizes were small.

Conclusion: Female sex and femininity predict symptom trajectories. No large sex differences in the strength of additional predictors were found, thus it may not be clinically useful to distinguish between predictors specific to male or female patients of persistent somatic symptoms.

Introduction

A substantial proportion of the general practitioner (GP) visits in the Netherlands (13%-43%) are related to common somatic symptoms for which no sufficient cause can be found after adequate physical examinations and interventions.[1,2] Persistence of these somatic symptoms is associated to increased functional impairment, feelings of internalized stigma and social isolation.[3] In addition to personal hardship, persistent common somatic symptoms may pose an economic burden on both an individual and societal level.[4,5]

A systematic review from 2009 showed that 10% to 30% of patients with medically unexplained somatic symptoms attending the GP or secondary care clinic did not improve during their follow-up period of 6 to 15 months.[6] More recent studies suggest higher rates of non-remission in primary care with 0.5 year (55.1%)[7], 1 year (51.2%)[8] and 2 years follow-up (56.8% and 37.1%).[9,10] However, studies describing the persistence of common somatic symptoms are difficult to compare due to methodological differences. Furthermore, these figures might not be representative of the general adult population. The prognosis of common somatic symptoms in the general population is likely to be more favorable, since by definition patient populations suffer from symptoms that they regard serious enough to visit a physician. To the best of our knowledge, only one recent study on the persistence of unexplained common somatic symptoms has been conducted in the adult general population, indicating that 36.4% of people had persistent common somatic symptoms measured over 3 years.[11]

A variety of predictors for persisting common somatic symptoms has been identified in adolescent and adult populations, including female sex,[8,12] physical and psychiatric comorbidities,[6] symptom characteristics (such as the duration, severity and heterogeneity),[9,13] and psychological traits (such as neuroticism, perfectionism and health perceptions).[12,14,15] Identification of predictors for persisting somatic symptoms is pivotal, as it allows for early detection, diagnosis and treatment and it could provide concrete starting points for interventions aiming to reduce or prevent such symptoms as well.[6] However, the definition of the persistence of symptoms in most aforementioned studies aiming to identify predictors was suboptimal. For example, persistence was defined based on an arbitrary number of contacts with the GP.[2] This does not distinguish between one's symptoms and healthcare-seeking behavior, which is especially problematic given the observation that patients who do not return to the GP often still experience symptoms.[16] Furthermore, it remains unknown whether these predictors differ between women and men.

Recently, a cross-sectional study found that not only female sex, but also feminine gender, which encompasses the roles, behaviors, identities and relationships of women prescribed by societal norms in a given context,[17] is associated with the severity of common somatic symptoms and prevalence of chronic diseases.[18] Gender, and its embodiment, is more dynamic than one's biological sex. Yet, gender and sex are often conflated in research. Therefore, to date it remains unknown whether sex and gender independently impact the severity and persistence of somatic symptoms in the general adult population. In addition, most studies assessing the severity and persistence of somatic symptoms have considered the cohort under study as a homogeneous population, while the longitudinal patterns of symptom development may show significant heterogeneity in their directions[10]. This means that in most studies variable patterns of somatic symptoms over time remain undetected.

We present the first large epidemiological cohort study to identify the predictors of longitudinal patterns of somatic symptoms in the general adult population, with a special emphasis on sex and gender differences. First, we will use latent class trajectory modeling to identify developmental patterns of symptom severity. Second, we will assess which predictors are associated with the different trajectories. Third, we aim to study whether the identified predictors for persistence of common somatic symptoms differ between females and males. We hypothesize that female sex and femininity associate with increased severity of common somatic symptoms.

Methods

Setting

This study is based on data collected within the Dutch Lifelines Cohort Study. The Lifelines Cohort Study is a multi-disciplinary prospective population-based cohort study examining in a unique three-generation design the health and health-related behaviors of 167,729 persons living in the North of The Netherlands. It employs a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioral, physical and psychological factors, which contribute to the health and disease of the general population, with a special focus on multi-morbidity and complex genetics. Extensive information on the cohort, design considerations and recruitment procedures is provided elsewhere.[19,20] The Lifelines Cohort Study is performed according to the principles of the Declaration of Helsinki and in accordance with the UMCG's research code. The Lifelines Cohort Study is approved by the Medical

Ethical Committee of the University Medical Center Groningen, The Netherlands[20]. For the current study, we adhered to the STROBE statement and GRoLTS guidelines for reporting of our findings.[21,22] To conform to the SAGER guidelines we reported our findings stratified by sex.[23]

Participants

Participants completed questionnaires on multiple topics including, but not limited to, demographics, health, personality, psychological and somatic symptoms, and psychosocial characteristics. These questionnaires asked for participants' biological sex. Hence, we refer to the participants as male and female, whereas we refer to masculinity and femininity when discussing gender.

In the current study, we used data from the adult participants gathered at four time points: at baseline (n=148,643; mean age 44.2 years [SD=12.8]; 58.6% females), at a first follow-up time point (n=124,443; mean age 46.5 years [SD=12.8]; 59.3% female; median time to follow-up: 13.0 [10-93] months), after a second follow-up time point (n=95,137; mean age 48.1 years [SD= 12.8]; 59.8% female; median time to follow-up: 25.0 [22-92] months), and after a third follow-up time point (n=90,077; mean age 49.8 years [SD=12.6]; 59.1% females; median time to follow-up: 46.0 [22-123] months). A more detailed overview of the population included is provided in ***Appendix A***. Attrition rates after 1.5, 3, and 4 years were 16.8%, 36.2% and 39.6%, respectively, compared to baseline. We did not find any indication for relevant systematic attrition: no meaningful associations between potential predictors of the severity of common somatic symptoms and attrition rates were found.

Variables

We assessed common somatic symptom severity by means of the 12-item ordinal Symptom CheckList-90 Somatization subscale (SCL-90 SOM; ***Appendix B***). The 12 items refer to how much bother or distress someone experienced the past week due to somatic symptoms. This scale has been recommended for large scale studies and has been shown to have sufficient measurement invariance over time, which makes it suitable to assess the measured concepts repeatedly over time.[24,25] The potential predictors for the persistence of common somatic symptoms, all assessed at baseline, are described in ***Table 1***. Femininity was operationalized via a recently developed gender index,[18] which accounts for the time-, place-, and society-bound nature of gender. In a subsample of adult Lifelines participants, with no suspected intersex condition or non-conform gender identity, a LASSO logistic regression model that predicts the participants' sex by means of psychosocial characteristics, including but not limited to hobbies, type of profession, dietary preferences and

time spent on household tasks, was calculated. In total, 85 unique psychosocial variables were included in the model (AUC=92%) and thus gendered. The included psychosocial predictors cover predominantly gender roles, and therefore reflect the gender roles as adhered to in the Lifelines cohort. The obtained estimates of the regression coefficients were applied to all adult Lifelines participants, providing each participant with an individual score on the gender index, i.e. participants' adherence to the gendered psychosocial variables. The gender index ranges from 0%, equaling masculinity, to 100%, equaling femininity. We assessed multicollinearity of the predictor variables by means of the variance inflation factor (VIF). We found no indication of problems with multicollinearity, as VIF was <5 in all analyses[26].

Statistical analyses

To identify different developmental trajectories of common somatic symptoms over time, latent class trajectory modeling was conducted in R version 3.5.2 and R studio 1.1.383 (R package *'lcmm'*, version 1.7.8).[27] Notably, these trajectories should not be reified, as these are merely estimated latent groups, not actual observed groups. An advantage of latent class trajectory modeling using full information maximum likelihood estimation as applied in this study is that it allows the number of times a participant was assessed to vary between participants, which facilitates the inclusion of participants with intermittent missing data or those who dropped out[27]. We used the GRoLTS guidelines and Lennon's *et al.* framework as a guidance to construct and interpret latent class trajectory modeling.[21,28]

Latent classes with different growth trajectories were modeled based on growth models that define how an outcome changes as a function of time, using an intercept and slope parameter. In order to find the model that best described the data, we fitted latent class growth models with fixed class-specific intercepts and slopes (LCGA), as well as more flexible growth mixture models (GMM) with (i) a random class-specific intercept and fixed slope per class and (ii) random class-specific intercepts and slopes. We fitted models with both linear and quadratic trajectories. LCGA and GMM models were fitted to the data with increasing number of classes (g=1 to g=7), after which indices of model fit were compared. Every model was run with multiple (25) random start values (derived from the one-class model) in order to identify a replicable Log-Likelihood maximum, that was unlikely to be at a local maximum. The best fitting model was then fully fitted with a maximum of 500 iterations. In the models the intercept and slope variances were constrained to be equal across classes. Data were rearranged as a function of chronological months since inclusion into the study. This resulted in 123 (baseline measurement being 0 months, the latest measurement being 123 months) instead of four assessment points, allowing for more complex

trajectories to be modeled. Data points at which no valid information on the SCL-90 SOM was provided were excluded (N=39 [0.02%], N=767 [0.62%], N=309 [0.33%] and N=320 [0.36%] participants at the first, second, third and fourth measurement, respectively). Ultimately, participants were allocated to a class based on their highest posterior class probability score. Participants with low posterior probabilities for all classes (<0.50) were excluded from the analyses.

The models with increasing number of classes were compared on four *a priori* formulated criteria: (i) the model with the lowest Bayesian Information Criterion (BIC) value was favorable[21,29]; (ii) the entropy of the model with the lowest BIC was assessed, as high entropy (>0.80) indicates strong distinctive capabilities between trajectory classes[30]; (iii) class sizes, as class sizes should not comprise less than 1% of the sample[31]; and (iv) theoretical plausibility, for example verifying whether the observed trajectories fit the longitudinal plots of raw data and previous empirical findings.[32]

To identify the predictors of somatic symptom trajectories, we conducted multiple logistic regression analyses, including all predictors as mentioned in ***Table 1*** as independent variables. To study whether the predictors for the latent subgroups differed between females and males we included interaction terms between sex and predictors.

To test whether the continuous covariates included in the multiple logistic regression analyses fulfilled the linearity assumption of multiple logistic regression we divided the covariates into quartiles, and assessed whether the estimates increased or decreased monotonically. IBM SPSS v. 25 was used to perform regression analyses. We maintained a two-sided alpha-value, corrected for multiple comparisons, of 0.001 (0.05/47, 24 predictors and 23 sex-by-predictor interaction terms within a family of tests).

Three sensitivity analyses were performed. First, we performed the regression analyses without adjusting for the presence of chronic diseases to explore its influence on the association between predictors and the identified trajectories. Second, we assessed whether the association between negative life events and the observed trajectories was partly explained by health-related negative life events. We excluded any health-related negative life events from the regression analyses to assess its influence on the association between negative life events and the identified trajectories. Third, we performed regression analyses with different symptom trajectories as a reference category.

Table 1. Overview of potential predictors for persistent common somatic symptoms (all assessed at baseline).

Potential predictor	Operationalization of predictor	Minimum - maximum in sample
Age	Municipally-registered age in years	18 – 92 years
Sex	Municipally-registered sex	a) Male b) Female
Gender	Gender Index (for an in-depth discussion of the construction of the gender index we refer to Ballering *et al.* (2020))	0% (reflecting masculinity) – 100% (reflecting femininity)
Educational level	Self-reported education[33]	a) High educational level b) Medium educational level c) Low educational level
Employment	Self-reported hours of paid work	0 – 60 hours
Chronic somatic diseases	Self-reported lifetime prevalence of chronic somatic diseases. A selection of diseases was included, based on (i) availability of data; (ii) the rankings of the Dutch Ministry of Health, Welfare and Sports of diseases causing the greatest loss of healthy life years in the adult general population and; (iii) previous literature.[11,34-36] Sex-specific diseases (e.g. prostate cancer) were excluded. Chronic somatic diseases included respiratory diseases (COPD, asthma), cardiovascular disorders (stroke, arrhythmia, heart attack, heart failures), intestinal disorders (Crohn's disease, ulcerative colitis, stomach ulcers), diabetes mellitus (type 1 and type 2), arthritis, osteoporosis, migraine, cancer (skin, colon and lung cancer), dementia, Parkinson's disease, kidney failure, epilepsy, rheumatoid arthritis and multiple sclerosis.	0 – 10 diseases
Physical functioning	Self-reported physical functioning (The mean score of the 10 items on the physical functioning subscale of the RAND-36)[37]	0 – 100 points
Emotional well-being	Self-reported emotional well-being (The mean score of the five items on the emotional well-being subscale of the RAND-36)[37]	0 – 100 points
Self-rated health	Self-reported health (The mean score of the five items on the general health perception subscale of the RAND-36)[37]	0 – 100 points

Table 1. Continued.

Potential predictor	Operationalization of predictor	Minimum - maximum in sample
Negative affect	Self-reported negative affect (The mean score of the 10 items on the negative affect subscale of the PANAS)[38]	0 – 4 points
Positive affect	Self-reported positive affect (The mean score of the 10 items on the positive affect subscale of the PANAS)[38]	0 – 4 points
Personality traits	Self-reported outcomes of the NEO PI R. Of 8 subscales of the NEO PI R, namely hostility, self-consciousness, impulsiveness, vulnerability, self-discipline, competence, deliberation and excitement the mean score was calculated. (Each subscale consists of 8 items)[39]	0 – 4 points
Negative life events	Self-reported presence of negative life events in the past year, based on the List of Threatening Experiences[40,41]	a) Any threatening experience present b) No threatening experience present
Longterm difficulties	Self-reported extent of long-term difficulties and stress, either somewhat or very much, based on the Longterm Difficulties Inventory[41]	a) Any longterm difficulty present b) No longterm difficulty present
Self-reported mood disorders	Self-reported lifetime prevalence of mood disorders (Mood disorders, based on DSM-IV, include major depression and bipolar disorder. No data on dysthymia was available.)	a) Any mood disorder present b) No mood disorder present
Self-reported anxiety disorders	Self-reported lifetime prevalence of anxiety disorders (Anxiety disorders, based on DSM-IV criteria, include panic disorder, agoraphobia, specific phobia and generalized anxiety disorder.)	a) Any anxiety disorder present b) No anxiety disorder present

Results

We found that the mean SCL-90 SOM score in the complete sample remained stable over time with scores of 1.36 (SD=0.38), 1.38 (SD=0.45), 1.42 (SD=0.44) and 1.36 (SD=0.42) at subsequent measurement waves.

Trajectory modeling

Of the fitted models, LCGA performed best (***Appendix C***). Therefore, only estimates of LCGA models are shown in ***Table 2***. The 5-class model fitted the data best, as is indicated by the lowest BIC value, good entropy and acceptable class sizes. The class-specific predicted mean SCL-90 SOM trajectories are displayed in ***Figure 1***. The first class, which comprises the majority of the population (N=113,444; 75.4%) reported minimal to no SCL-90 SOM symptoms over time. The second class (N=1,717; 1.1%) reported a high, stable SCL-90 SOM symptom score over time. The third class (N=7,168; 4.8%) showed slightly decreasing, intermediate SCL-90 SOM symptom score, whereas the fourth class (N=25,954; 17.3%) showed a low, stable SCL-90 SOM symptom score, albeit somewhat higher than the first class. The fifth class (N=2,211; 1.5%) started with a relatively low SCL-90 SOM score, which steeply increased over time. ***Appendix D*** shows plots with individual SCL-90 SOM score trajectories, stratified per class.

Logistic regression analyses

Participants were allocated to one of the trajectory classes, based on their posterior class probability score; 1,495 (1.0%) participants with low posterior probabilities for all classes (<0.50) were excluded from the analyses.

First, we identified the predictors of high, stable symptom severity (class 2) compared with low, stable symptom severity (class 4) by multiple logistic regression analyses (***Table 3***). Class 4 was selected as the reference category to facilitate a comparison with the subsequent analyses. We found that female sex was significantly associated with high symptom severity (OR=1.74; 95% CI=1.36-2.22), while femininity was associated with low symptom severity (OR=0.60; 95%CI= 0.44-0.82). Also, increased physical functioning and emotional wellbeing were associated with low symptom severity (OR=0.96; 95%CI=0.96-0.97 for both predictors). On the other hand, better self-rated health was associated with high symptom severity (OR=1.03; 95%CI=1.02-1.04). We found that personality traits were not statistically significantly associated with high symptom severity. We assessed the statistical significance of the interaction term between sex and all predictors. Only the interaction terms between sex and hours of paid employment

(OR=0.98; 95%CI=0.98-0.99), and sex and physical functioning (OR=0.99; 95%CI=0.98-0.99) were statistically significant, indicating that the negative association between these predictors and a high symptom severity was stronger in females than in males.

Second, we assessed which predictors were associated with increasing symptom severity, by multiple logistic regression analyses predicting an increasing SCL-90 SOM score over time (class 5) versus low, stable SCL-90 SOM score (class 4) as these trajectories had a similar intercept. This allows for identification of predictors that may associate with an increasing symptom severity, instead of a low, stable symptom severity. ***Table 4*** shows that females have 1.15 times the odds (95%CI=0.99-1.40) compared to males of having increasing symptom severity over time, however this result did not reach statistical significance. Femininity seemed to be protective of increasing symptom severity over time (OR=0.66; 95%CI=0.51-0.85), yet experiencing a negative life event is disadvantageous (OR=1.30; 95%CI=1.16-1.47). The OR of the interaction term between sex and education (OR=1.23; 95%CI=1.01-1.58), sex and the presence of chronic disease (OR=0.79; 95%CI=0.63-0.98), sex and physical functioning (OR=0.99; 95%CI=0.98-0.99), and sex and the score on the NEO-PI-R deliberation subscale (OR=1.35; 95%CI=1.10-1.64) were statistically significant, indicating that the association between these predictors and increasing symptom severity differed between females and males. The results of the sensitivity analyses, which assess the effect of the presence of chronic diseases and health-related negative life events on the association between femininity or sex and symptom trajectories, as well as the effect of selecting class 1 (no, stable symptoms) as the reference symptom trajectory instead of class 4 (low, stable symptoms) yielded essentially the same results and are shown in ***Appendix E***.

Table 2: Parameter estimates for 1 to 7 classes (N=150,494) using a linear trajectory function.

G	NPM[a]	Log-Likelihood	BIC[b] value	Entropy
1	3	-254288	508611	100%
2	6	-181449	362969	91.6%
3	9	-161083	322273	87.6%
4	12	-154669	309482	85.6%
5[c]	15	-149292	298764	86.1%
6	18	-154669	309553	51.8%
7	21	n.a.	n.a.	n.a.

[a]Number of parameters; [b]Bayesian Information Criteria; [c]Preferred model.

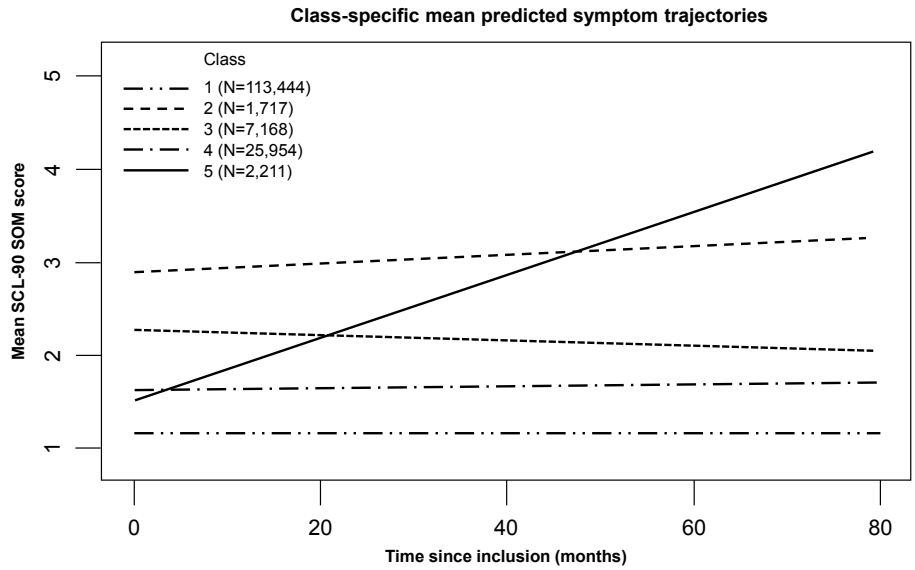


Figure 1. Class-specific mean predicted symptom trajectories.

Participants per class (%)						Mean posterior class probability score
150,494 (100%)						1
132,523 (88.1%)	17,971 (11.9%)					0.98/0.92
24,711 (16.4%)	5,516 (6.7%)	120,267 (79.9%)				0.86/0.92/0.96
1,985 (1.3%)	8,054 (5.4%)	114,618 (76.2%)	25,837 (17.2%)			0.91/0.86/0.94/0.81
113,444 (75.4%)	1,717 (1.1%)	7,168 (4.8%)	25,954 (17.3%)	2,211 (1.5%)		0.94/0.90/0.85/0.81/0.82
111,525 (74.1%)	8,054 (5.4%)	28,930 (19.2%)	1,985 (1.3%)	0 (0%)	0 (0%)	0.61/0.86/0.77/0.91/0/0
n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	Failed to converge

Table 3: The associations between predictors and high, stable symptom severity over time.

		Odds ratio (95% CI)[b]		
Predictors		**Total (N=27,183)**	**Males (N=8,084)**	**Females (N=19,099)**
Sex	Male	1	n.a.	n.a.
	Female	1.74 (1.36-2.22)*	n.a.	n.a.
Femininity		0.60 (0.44-0.82)*	0.65 (0.39-1.08)	0.56 (0.37-0.85)*
Age		0.99 (0.98-1.00)	0.98 (0.97-1.00)	0.99 (0.98-1.00)
Education	Low	1	1	1
	Medium	0.75 (0.65-0.87)*	0.81 (0.61-1.08)	0.74 (0.62-0.88)*
	High	0.46 (0.37-0.57)*	0.43 (0.28-0.67)*	0.49 (0.38-0.63)*
Hours of paid employment[a]		0.99 (0.99-1.00)	1.00 (1.00-1.01)	0.99 (0.98-0.99)*
Presence of chronic disease		1.55 (1.33-1.81)*	1.74 (1.30-2.32)*	1.47 (1.22-1.77)*
Physical functioning[a]		0.96 (0.96-0.97)*	0.97 (0.96-0.97)*	0.96 (0.95-0.96)*
Emotional wellbeing		0.96 (0.96-0.97)*	0.96 (0.95-0.97)*	0.96 (0.96-0.97)*
Self-rated health		1.03 (1.02-1.04)*	1.04 (1.03-1.05)*	1.03 (1.03-1.04)*
Negative affect		2.20 (1.87-2.58)*	2.12 (1.56-2.86)*	2.21 (1.83-2.66)*
Positive affect		1.34 (1.14-1.57)*	1.10 (0.81-1.49)	1.35 (1.03-1.77)*
NEO PI R	Anger	1.14 (0.99–1.32)	0.91 (0.63-1.30)	1.16 (0.98-1.38)
	Self-consciousness	0.81 (0.70–0.92)*	1.14 (0.87-1.49)	0.79 (0.68-0.93)*
	Impulsivity	0.92 (0.78-1.08)	1.15 (0.81-1.62)	0.87 (0.73-1.05)
	Vulnerability	0.75 (0.63-0.90)*	0.61 (0.43-0.89)*	0.81 (0.66-1.00)
	Self-discipline	1.12 (0.96-1.30)	1.16 (0.87-1.56)	1.11 (0.93-1.33)
	Competence	0.91 (0.73-1.13)	0.99 (0.63-1.53)	0.88 (0.68-1.14)
	Deliberation	0.87 (0.75-1.01)	0.84 (0.62-1.13)	0.88 (0.74-1.06)
	Excitement	0.97 (0.85-1.10)	0.96 (0.74-1.25)	0.96 (0.82-1.12)
Occurrence of negative life event		1.17 (0.98-1.39)	1.39 (0.97-1.79)	1.11 (0.91-1.36)
Occurrence of long term difficulty		1.23 (0.88-1.70)	0.86 (0.46-1.62)	1.38 (0.94-2.02)
Presence of mood disorder		1.05 (0.90-1.23)	1.13 (0.82-1.55)	1.03 (0.86-1.23)
Presence of anxiety disorder		1.27 (1.06-1.52)*	1.01 (0.69-1.47)	1.36 (1.11-1.67)*

[a]Interaction terms between these predictors and sex were statistically significant. [b]Please note that the odds presented are per unit change on the scale of the predictor, thus magnitudes are not always directly comparable. *Indicates statistical significance (p<0.001) Note: Nagelkerke's R^2 for the model including all participants, only the men and only the women allocated to class 2 and class 4 are 0.38, 0.36 and 0.39, respectively.

Table 4. The associations between multiple predictors and increasing symptom severity over time.

			Odds ratio (95% CI)[b]	
Predictors		**Total (N=27,572)**	**Males (N=8,288)**	**Females (N=19,284)**
Sex	Male	1	n.a.	n.a.
	Female	1.15 (0.99-1.40)	n.a.	n.a.
Femininity		0.66 (0.51-0.85)*	0.83 (0.55-1.24)	0.61 (0.44-0.85)*
Age		1.00 (1.00-1.01)	1.01 (1.00-1.02)	1.00 (0.99-1.01)
Education	Low	1	1	1
	Medium[a]	0.87 (0.77-0.98)*	0.71 (0.57-0.88)*	0.95 (0.83-1.01)
	High	0.71 (0.61-0.82)*	0.62 (0.47-0.80)*	0.77 (0.64-0.93)*
Hours of paid employment		1.00 (0.99-1.00)	1.00 (0.99-1.01)	0.99 (0.99-1.00)
Presence of chronic disease[a]		0.98 (0.88-1.09)	1.18 (1.01-1.42)*	0.89 (0.79-1.02)
Physical functioning[a]		1.00 (1.00-1.01)	1.01 (1.00-1.01)	1.00 (0.99-1.00)
Emotional wellbeing		1.01 (1.01-1.02)*	1.01 (1.00-1.02)	1.01 (1.00-1.02)
Self-rated health		1.00 (1.00-1.01)	1.00 (0.99-1.01)	1.00 (1.00-1.01)
Negative affect		1.12 (0.99-1.28)	0.98 (0.77-1.25)	1.18 (1.01-1.39)*
Positive affect		1.00 (0.87-1.15)	0.83 (0.65-1.06)	1.09 (0.92-1.30)
NEO PI R	Anger	1.12 (1.00-1.26)	1.23 (1.01-1.51)*	1.07 (0.93-1.24)
	Self-consciousness	1.01 (0.90-1.12)	0.99 (0.80-1.23)	1.01 (0.88-1.16)
	Impulsivity	1.03 (0.91-1.18)	1.13 (0.88-1.44)	1.00 (0.86-1.17)
	Vulnerability	0.96 (0.83-1.12)	0.88 (0.66-1.17)	0.99 (0.82-1.18)
	Self-discipline	1.17 (1.03-1.33)*	1.26 (1.00-1.58)	1.12 (0.96-1.29)
	Competence	0.89 (0.74-1.07)	0.85 (0.61-1.18)	0.90 (0.72-1.12)
	Deliberation[a]	1.02 (0.90-1.16)	0.83 (0.67-1.04)	1.11 (0.95-1.29)
	Excitement	0.97 (0.88-1.08)	1.05 (0.87-1.27)	0.94 (0.83-1.07)
Occurrence of negative life event		1.30 (1.16-1.47)*	1.11 (0.91-1.37)	1.40 (1.21-1.62)*
Occurrence of long term difficulty		0.96 (0.81-1.13)	0.96 (0.71-1.28)	0.97 (0.79-1.18)
Presence of mood disorder		1.11 (0.99-1.31)	1.23 (0.93-1.62)	1.11 (0.94-1.30)
Presence of anxiety disorder		1.05 (0.88-1.26)	1.46 (1.04-2.04)*	0.95 (0.77-1.18)

[a]Interaction terms between these predictors and sex were statistically significant. [b]Please note that the odds presented are per unit change on the scale of the predictor, thus magnitudes are not always directly comparable. *Indicates statistical significance (p<0.001). Nagelkerke's R^2 for the model including all participants, only the men and only the women allocated to class 5 and class 4 are 0.11, 0.28 and 0.11, respectively.

Discussion

To the best of our knowledge, this is the first large general population cohort study that assesses longitudinal somatic symptom trajectories by means of LCGA. This data-driven method allows for the identification of homogeneous patterns of symptom severity over time from heterogeneous data. We found that a five-class linear model that excluded intraclass individual variation fitted the data best. The majority of the cohort had a stable symptom trajectory (93.7%), with low (class 1; 75.4%), slightly higher (class 4; 17.3%) and high (class 2; 1.1%) symptom severity. In addition, we identified a class with slightly decreasing (class 3; 4.8%) and a class with increasing (class 5; 1.5%) symptom severity over time. We found that female sex is a predictor for a high, stable SCL-90 SOM score. However, female sex only approached statistical significance for an increasing SCL-90 SOM score, compared to male sex. Femininity, in contrast, appeared to be protective for both a stable and an increasing somatic symptom severity. In females, hours of paid employment and physical functioning were more strongly negatively associated with stable symptom severity than in males. Regarding increasing symptom severity over time, education, the presence of chronic disease, physical functioning and the score on the NEO-PI-R deliberation subscale differed in predictive strength between females and males.

Strengths and limitations

The principal strength of this study is the data-driven approach we used to estimate common somatic symptom trajectories in a large general population cohort. Latent class trajectory modeling allows for identifying nuances between seemingly similar subpopulations.[28] This inductive approach facilitates the identification of novel predictors of at risk subpopulations, especially if individuals' symptom trajectories are analyzed as an outcome, rather than as an independent variable. However, note that when analysis involves latent class trajectories either as outcome or exposure, one should not view the trajectories as concrete entities, but rather as a method to reduce the observed heterogeneity in the data[42]. Furthermore, our cohort had a large sample size and was followed up for a long period of time, with multiple measurements, allowing for complex models to be fitted. Lastly, the incorporation of participants' gender is advantageous, as it allows for disentangling the biological and psychosocial influences related to being a woman or man on somatic symptom trajectories.

Our study had several limitations. First, we assessed the mean SCL-90 SOM score as an aggregate score and therefore we have not differentiated between individual symptoms. Possibly, participants had different symptoms that bothered them over time, despite their symptom scores remaining stable. Second, we could not account

for illness cognitions or health care utilization, as no data hereon were available. Illness cognitions are thought to account for 30%-40% of the variance in health outcomes related to somatic symptoms,[43] such as the persistence of symptoms.[44] Similarly, health care utilization is known to associate somatic symptom burden, and thus may affect symptom trajectories.[45] Also, we only assessed predictors at baseline, but predictors may also have an influence during the course of one's symptoms, such as the development of chronic diseases or health care utilization. All predictors are self-reported, which means that the measures of the life-time prevalence of mood and anxiety disorders are not necessarily a clinical diagnosis, and the latter two may be more reflective of experienced mood and anxiety symptoms than of a clinical diagnosis.

Latent class trajectories

We identified three stable symptom severity trajectories (93.7%) and two relatively small classes that follow a consistently increasing (1.5%) and decreasing (4.8%) course. The proportion of participants with non-stable trajectories is in line with previous research that used latent class trajectory modeling. In a patient population with medically unexplained somatic symptoms 92.6% of the patients had a stable symptom score over the two-years follow-up. The remaining 7.4% of patients improved. However, this study used the Patient Health Questionnaire-15 to measure symptom severity, which differs from the current SCL-90 SOM subscale.[10] Another study conducted in a general adolescent cohort showed that 85.3% of adolescents had a predominantly stable symptom trajectory over time.[12] Four trajectories were identified in this study. Again, this study used a different questionnaire, but more importantly, an adolescent cohort might not be directly comparable to an adult cohort. It is disputed whether somatic symptoms in adolescents and adults are comparable in onset, healthcare-seeking behavior and treatment, possibly due to stronger family influences in adolescents and a differing physiology compared to adults.[46] Furthermore, adolescence is thought to be accompanied by a heightened bodily awareness and therefore with the experience of common somatic symptoms.[47] Overall, stable somatic symptom scores over time prevail in the aforementioned studies despite the differing study populations.

Sex and gender in relation to somatic symptom trajectories

In line with our current study, recent studies have found female sex to be associated with more numerous and more severe somatic symptoms,[18,48] as well as with an increasing severity of somatic symptoms over time.[12,15] Multiple explanations have been raised for this phenomenon. First, females may have a heightened pain sensitivity due to biological differences.[49] Sex hormones, genotypes, immune systems and neurology

may induce differences in the processing of pain that predispose females to worse symptom trajectories than males.[50] Second, females are thought to be more aware of bodily sensations than males. This heightened awareness allows for easier and earlier perception of somatic symptoms in females than in males.[51] These biological differences may explain the female preponderance in somatic symptoms as found in our study, but our results also point toward a role for psychosocial gender differences.

Femininity was found to be protective against both a high and increasing symptom severity. This is different from earlier cross-sectional studies that showed an association between femininity, measured by the gender index, domestic responsibilities or the BEM sex role inventory, respectively, and higher levels of common somatic symptoms[18,52] or that found no association.[53] These differences may be explained by the longitudinal nature of the current study, which provides insight into the dynamics of symptoms over time and may result in a more precise assessment of somatic symptom severity. Furthermore, in the former study in which the gender index was used, all adult Lifelines participants were included and femininity was found to be associated with more severe symptoms. However, the current study focusses on participants with high or increasing symptom severity (1.1% and 1.5% of the adult participants, respectively) and femininity was not found to associate with increased symptom severity. Possibly increased healthcare-seeking behavior plays a role in femininity being a protective factor for both high, stable and increasing symptom severity over time,[8] as healthcare-seeking behavior is known to be gendered[54] and may prevent worsening of symptoms over time.[45] Feminine people are thought to have a lower threshold to seek help or medical care, especially from their GP.[55] Femininity is, for example, related to providing and facilitating care for the family, allowing feminine people to be more often in contact with healthcare providers, concomitantly lowering the barrier for healthcare-seeking behavior. Additionally, femininity is related to being open and less stoic about one's symptoms, facilitating healthcare-seeking behavior. Masculinity in contrast, relates to being less expressive about distress and seeking help for symptoms is stereotypically seen as socially undermining an individual's masculinity.[51,56] The earlier study that used the gender index included participants with low symptoms severity as well, in which healthcare-seeking behavior may not be as important.

In addition, what constitutes femininity differs between studies and changes over time and place, yielding different results. For example, the BEM sex role inventory was developed in 1974 and then widely applied, but is currently deemed to hold limited validity as an operationalization of femininity or masculinity.[57] Lastly, the association between femininity and higher levels of somatic symptom severity in the previous

studies may have been partially explained by the presence of chronic diseases, whilst we adjusted for this in the current study. Sensitivity analyses showed that indeed an adjustment for the presence of chronic diseases slightly strengthened the protective association between femininity and high or increasing symptom severity over time.

Predictors of stable and increasing common somatic symptoms

In addition to sex and gender, multiple factors were predictive of persistent common somatic symptoms. Higher education, higher levels of physical functioning and higher emotional wellbeing at baseline are associated with low, stable symptom severity. This is in line with previous research that suggests that these factors have a positive influence on overall functioning and wellbeing.[11] Here, we also found that the lower one rates his or her own general health, the lower the odds that one has persisting common somatic symptoms. Perceptions of low general health may prompt one to seek medical help, which may lead to an improvement of symptoms.[8] However, previous studies contradict each other with regards to self-rated health and the course of somatic symptoms.[8,11,12] These contradictions might be due to the different conceptualizations of self-related health and somatic symptoms in the studies.

We also found that the presence of anxiety disorders is related to stable and increasing symptom severity in females. Anxiety disorders are diagnosed approximately twice as often in females than in males,[58] and often manifest themselves with prominent somatic characteristics[59]. Therefore, the presence of anxiety disorders may contribute to the elevated SCL-90 SOM scores as found in this study. A self-reported mood disorder, however, was not associated with a high or increasing symptom severity. Evidence from cross-sectional studies suggests that mood disorders are associated with more severe common somatic symptoms.[59,60] In contrast, results of longitudinal research assessing mood disorders in relation to somatic symptoms are contradictive.[61] To date, it has not been possible to draw any definitive conclusion on whether mood disorders predict an unfavorable somatic symptom prognosis. It has been argued that a similar mechanism as mentioned above may apply to people with mood disorders: being affected by mood disorders may prompt healthcare-seeking behavior, resulting in an improvement of symptoms. The aforementioned longitudinal studies, however, including our study, do not differentiate between, or assess different, somatic symptoms. Thus the association between mood disorders and one type of symptoms may be overshadowed by the lack of an association with other types of symptoms.

We also found that negative life events are predictors of increasing symptom severity. As a sensitivity analysis, we removed any item from the negative life events scale that was related to experiencing a severe disease. The direction and strength of

the association remained similar. It is thought that psychological distress as a consequence of a negative life event may result in somatic symptoms.[62] Physiological and emotional stress-mechanisms are suggested as the link between psychological distress and somatic symptoms.[63] Such mechanisms heighten one's bodily vigilance, consequently facilitating people interpret bodily signals more easily as somatic symptoms.

Implications for further research and clinical practice

Further research could focus on symptom-specific latent class trajectories, to assess whether differences in trajectories exist between symptoms. Additionally, one could study whether the type of reported symptoms change over time in the classes with stable, high mean symptom severity scores and whether these changes follow specific sequences. The results from this study also show that the majority of the general population remains stable in their level of symptom severity and that only a relatively small proportion has a high or increasing symptom severity. However, it remains unknown whether it is merely the latter population that seeks medical attention and if so, what factors are associated with this healthcare-seeking behavior. As a protective association between femininity and a high and increasing symptom severity was found in this study, it is especially interesting to study to what extent femininity relates to healthcare-seeking behavior for somatic symptoms. For those patients who visit their GP, it is pivotal that predictors for increasing symptom severity are recognized, preferably in an early stage.

We found no large sex differences in the predictors of high or increasing symptom severity, thus it may not be clinically useful to distinguish between predictors specific to male or female patients with persistent common somatic symptoms. Furthermore, for reasons of clarity we currently described the associations of sex and gender with common somatic symptom trajectories separately. However, although sex and gender are different concepts, a clear demarcation between these in clinical practice is artificial: clinicians cannot consider sex without gender and vice versa as these concepts are intertwined.

References

1. Olde Hartman TC, Blankenstein AH, Molenaar AO, et al. NHG-standaard somatisch onvoldoende verklaarde lichamelijke klachten (SOLK). *Huisarts Wet*. 2013;56(5):222-230.
2. Verhaak PF, Meijer SA, Visser AP, Wolters G. Persistent presentation of medically unexplained symptoms in general practice. *Family practice*. 2006;23(4):414-420.
3. Dirkzwager AJ, Verhaak PF. Patients with persistent medically unexplained symptoms in general practice: Characteristics and quality of care. *BMC Family Practice*. 2007;8(1):33.
4. Konnopka A, Schaefert R, Heinrich S, et al. Economics of medically unexplained symptoms: A systematic review of the literature. *Psychotherapy and Psychosomatics*. 2012;81(5):265-275. https://www.karger.com/DOI/10.1159/000337349. doi: 10.1159/000337349.
5. Joustra ML, Janssens KAM, Bültmann U, Rosmalen JGM. Functional limitations in functional somatic syndromes and well-defined medical diseases. results from the general population cohort LifeLines. *Journal of Psychosomatic Research*. 2015;79(2):94-99. doi: 10.1016/j.jpsychores.2015.05.004.
6. olde Hartman TC, Borghuis MS, Lucassen PL, van de Laar FA, Speckens AE, van Weel C. Medically unexplained symptoms, somatisation disorder and hypochondriasis: Course and prognosis. A systematic review. *Journal of Psychosomatic Research*. 2009;66(5):363-377. doi: 10.1016/j.jpsychores.2008.09.018.
7. Lamahewa K, Buszewicz M, Walters K, Marston L, Nazareth I. Persistent unexplained physical symptoms: A prospective longitudinal cohort study in UK primary care. *British journal of general practice*. 2019;69(681):e246-e253. doi: 10.3399/bjgp19X701249.
8. Steinbrecher N, Hiller W. Course and prediction of somatoform disorder and medically unexplained symptoms in primary care. *General Hospital Psychiatry*. 2011;33(4):318-326. doi: 10.1016/j.genhosppsych.2011.05.002.
9. Budtz-Lilly A, Vestergaard M, Fink P, Carlsen AH, Rosendal M. The prognosis of bodily distress syndrome: A cohort study in primary care. *General Hospital Psychiatry*. 2015;37(6):560-566. doi: 10.1016/j.genhosppsych.2015.08.002.
10. Claassen - van Dessel N, van der Wouden JC, Hoekstra T, Dekker J, van der Horst HE. The 2-year course of medically unexplained physical symptoms (MUPS) in terms of symptom severity and functional status: Results of the PROSPECTS cohort study. *Journal of Psychosomatic Research*. 2018;104:76-87. doi: 10.1016/j.jpsychores.2017.11.012.
11. van Eck van der Sluijs, J. F., ten Have M, de Graaf R, Rijnders CAT, van Marwijk HWJ, van der Feltz-Cornelis CM. Predictors of persistent medically unexplained physical symptoms: Findings from a general population study. *Frontiers in Psychiatry*. 2018;9:613. doi: 10.3389/fpsyt.2018.00613.
12. Janssens KAM, Klis S, Kingma EM, Oldehinkel AJ, Rosmalen JGM. Predictors for persistence of functional somatic symptoms in adolescents. *The Journal of pediatrics*. 2014;164(4):900-905(e2). doi: rg/10.1016/j.jpeds.2013.12.003.
13. Kooiman CG, Bolk JH, Rooijmans HGM, Trijsburg RW. Alexithymia does not predict the persistence of medically unexplained physical symptoms. *Psychosomatic medicine*. 2004;66(2):224-232.
14. Bonvanie IJ, Rosmalen JGM, van Rhede van der Kloot, C. M., Oldehinkel AJ, Janssens KAM. Functional somatic symptoms are associated with perfectionism in adolescents. *J Psychosom Res*. 2015;79(4):328-330.

15. De Gucht V, Fischler BF, Heiser W. Personality and affect as determinants of medically unexplained symptoms in primary care; A follow-up study. *Journal of Psychosomatic Research*. 2004;56(3):279-285. doi: 10.1016/S0022-3999(03)00127-2.
16. Koch H, van Bokhoven MA, Bindels P, van der Weijden T, Dinant GJ, ter Riet G. The course of newly presented unexplained complaints in general practice patients: A prospective cohort study. *Family practice*. 2009;26(6):455-465. https://doi.org/10.1093/fampra/cmp067. doi: 10.1093/fampra/cmp067.
17. Johnson JL, Greaves L, Repta R. Better science with sex and gender: Facilitating the use of a sex and gender-based analysis in health research. *International Journal for Equity in Health*. 2009;8(1):14. doi: 10.1186/1475-9276-8-14.
18. Ballering AV, Bonvanie IJ, Olde Hartman TC, Monden R, Rosmalen JGM. Gender and sex independently associate with common somatic symptoms and lifetime prevalence of chronic disease. *Social Science & Medicine*. 2020;253:112968. doi: 10.1016/j.socscimed.2020.112968.
19. Klijs B, Scholtens S, Mandemakers JJ, Snieder H, Stolk RP, Smidt N. Representativeness of the LifeLines cohort study. *PLoS ONE*. 2015;10(9):e0137203. doi: 10.1371/journal.pone.0137203.
20. Scholtens S, Smidt N, Swertz MA, et al. Cohort profile: LifeLines, a three-generation cohort study and biobank. *International Journal of Epidemiology*. 2015;44(4):1172-80. doi: 10.1093/ije/dyu229.
21. Van De Schoot R, Sijbrandij M, Winter SD, Depaoli S, Vermunt JK. The GRoLTS-checklist: Guidelines for reporting on latent trajectory studies. *Structural Equation Modeling: A Multidisciplinary Journal*. 2017;24(3):451-467. doi: 10.1080/10705511.2016.1247646.
22. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. *International Journal of Surgery*. 2014;12(12):1495-1499. doi: 10.1016/j.ijsu.2014.07.013.
23. Heidari S, Babor TF, De Castro P, Tort S, Curno M. Sex and gender equity in research: Rationale for the SAGER guidelines and recommended use. *Research Integrity and Peer Review*. 2016;1:2. doi: 10.1186/s41073-016-0007-6.
24. Rytilä-Manninen M, Fröjd S, Haravuori H, et al. Psychometric properties of the symptom checklist-90 in adolescent psychiatric inpatients and age- and gender-matched community youth. *Child and Adolescent Psychiatry and Mental Health*. 2016;10(23):23. doi: 10.1186/s13034-016-0111-x.
25. Zijlema WL, Stolk RP, Löwe B, Rief W, White PD, Rosmalen JGM. How to assess common somatic symptoms in large-scale studies: A systematic review of questionnaires. *Journal of Psychosomatic Research*. 2013;74(6):459-68. doi: 10.1016/j.jpsychores.2013.03.093.
26. Miles J, Shevlin M. Collinearity. In: *Applying regression and correlation: A guide for students and researchers.* Sage; 2001:126-135.
27. Proust-Lima C, Philipps V, Liquet B. Estimation of extended mixed models using latent classes and latent processes: The R package lcmm. *Journal of Statistical Software*. 2017;78(2). https://www.jstatsoft.org/v078/i02. doi: 10.18637/jss.v078.i02.
28. Lennon H, Kelly S, Sperrin M, et al. Framework to construct and interpret latent class trajectory modelling. *BMJ Open*. 2018;8(7):e020683-020683.
29. Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: A monte carlo simulation study. *Structural equation modeling: A multidisciplinary Journal*. 2007;14(4):535-569. doi: 10.1080/10705510701575396.
30. Celeux G, Soromenho G. An entropy criterion for assessing the number of clusters in a mixture model. *Journal of classification*. 1996;13(2):195-212.

31. Infurna FJ, Grimm KJ. The use of growth mixture modeling for studying resilience to major life stressors in adulthood and old age: Lessons for class size and identification and model selection. *The Journals of Gerontology: Series B*. 2017;73(1):148-159.
32. Ram N, Grimm KJ. Methods and measures: Growth mixture modeling: A method for identifying differences in longitudinal change among unobserved groups. *International journal of behavioral development*. 2009;33(6):565-576. doi: 10.1177/0165025409343765.
33. van Zon SKR, Snieder H, Bültmann U, Reijneveld SA. The interaction of socioeconomic position and type 2 diabetes mellitus family history: A cross-sectional analysis of the lifelines cohort and biobank study. *BMJ Open*. 2017;7(4):e015275. doi: 10.1136/bmjopen-2016-015275.
34. Dutch Ministry of Health, Welfare and Sports. Ranglijst aandoeningen op basis van verlies aan gezonde levensjaren (ziektejaarequivalenten). https://www.volksgezondheidenzorg.info/ranglijst/ranglijst-aandoeningen-op-basis-van-verlies-aan-gezonde-levensjaren-ziektejaarequivalenten. Updated 2019. Accessed May 23rd, 2019.
35. Hoeymans N, Gijsen R, Slobbe LC. 59 important health problems; a selection of diseases for public health monitoring. *Nederlands tijdschrift voor geneeskunde*. 2013;157(31):A5994.
36. Gijsen R, Poos MJJC, Slobbe LC, Mulder M, In 't Panhuis-Plasmans MHD, Hoeymans N. Een nieuwe selectie van ziekten voor de volksgezondheid toekomst verkenningen. *RIVM*. 2013;010003004/2013.
37. Aaronson NK, Muller M, Cohen PDA, et al. Translation, validation, and norming of the dutch language version of the SF-36 health survey in community and chronic disease populations. *J Clin Epidemiol*. 1998;51(11):1055-1068. doi: 10.1016/S0895-4356(98)00097-3.
38. Hill RD, Van Boxtel M, Ponds R, Houx PJ, Jolles J. Positive affect and its relationship to free recall memory performance in a sample of older dutch adults from the maastricht aging study. *International Journal of Geriatric Psychiatry*. 2005;20(5):429-435. doi: 10.1002/gps.1300.
39. Costa PT, McCrae RR. Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI) professional manual. *Psychological Assessment Resources*. 1992.
40. Brugha TS, Cragg D. The list of threatening experiences: The reliability and validity of a brief life events questionnaire. *Acta Psychiatr Scand*. 1990;82(1):77-81.
41. Rosmalen JGM, Bos EH, de Jonge P. Validation of the long-term difficulties inventory (LDI) and the list of threatening experiences (LTE) as measures of stress in epidemiological population-based cohort studies. *Psychological Medicine*. 2012;42(12):2599-608. doi: 10.1017/S0033291712000608.
42. van Loo HM, Wanders RB, Wardenaar KJ, Fried EI. Problems with latent class analysis to detect data-driven subtypes of depression. *Molecular psychiatry*. 2018;23(3):495-496. doi: 10.1038/mp.2016.202.
43. McAndrew LM, Crede M, Maestro K, Slotkin S, Kimber J, Phillips LA. Using the common-sense model to understand health outcomes for medically unexplained symptoms: A meta-analysis. *Health psychology review*. 2019;13(4):427-446.
44. Moss-Morris R, Spence MJ, Hou R. The pathway from glandular fever to chronic fatigue syndrome: Can the cognitive behavioural model provide the map? *Psychological Medicine*. 2011;41(5):1099-1107.
45. Lee S, Creed FH, Ma Y, Leung CM. Somatic symptom burden and health anxiety in the population and their correlates. *J Psychosom Res*. 2015;78(1):71-76. doi: 10.1016/j.jpsychores.2014.11.012.
46. Weisblatt E, Hindley P, Rask CU. Medically unexplained symptoms in children and adolescents. In: Creed F, Henningsen P, Fink P, eds. *Medically unexplained symptoms, somatisation and bodily distress - developing better clinical services.* 1st ed. Cambridge, UK: Cambridge University Press; 2011:158-174.

47. Rhee H. Relationships between physical symptoms and pubertal development. *Journal of Pediatric Health Care*. 2005;19(2):95-103.
48. Tomenson B, Essau C, Jacobi F, et al. Total somatic symptom score as a predictor of health outcome in somatic symptom disorders. *The British Journal of Psychiatry*. 2013;203(5):373-380. doi: 10.1192/bjp.bp.112.114405.
49. Fillingim RB. Sex, gender, and pain. In: Legato MJ, ed. *Principles of gender-specific medicine*. 3rd ed. San Diego: Academic Press; 2017:481-496. 10.1016/B978-0-12-803506-1.00038-3.
50. Bartley EJ, Fillingim RB. Sex differences in pain: A brief review of clinical and experimental findings. *Br J Anaesth*. 2013;111(1):52-58. doi: 10.1093/bja/aet127.
51. Barsky AJ, Peekna HM, Borus JF. Somatic symptom reporting in women and men. *Journal of General Internal Medicine*. 2001;16(4):266-275. doi: 10.1046/j.1525-1497.2001.00229.x.
52. Krantz G, Ostergren P. Double exposure: The combined impact of domestic responsibilities and job strain on common symptoms in employed swedish women. *The European Journal of Public Health*. 2001;11(4):413-419.
53. Castro Y, Carbonell JL, Anestis JC. The influence of gender role on the prediction of antisocial behaviour and somatization. *Int J Soc Psychiatry*. 2012;58(4):409-416. doi: 10.1177/0020764011406807.
54. Hart CG, Saperstein A, Magliozzi D, Westbrook L. Gender and health: Beyond binary categorical measurement. *J Health Soc Behav*. 2019;60(1):101-118.
55. Loikas D, Karlsson L, von Euler M, Hallgren K, Schenck-Gustafsson K, Bastholm Rahmner P. Does patient's sex influence treatment in primary care? experiences and expressed knowledge among physicians--a qualitative study. *BMC family practice*. 2015;16:137. doi: 10.1186/s12875-015-0351-5.
56. MacLean A, Sweeting H, Hunt K. 'Rules' for boys,'guidelines' for girls: Gender differences in symptom reporting during childhood and adolescence. *Social Science & Medicine*. 2010;70(4):597-604. doi: 10.1016/j.socscimed.2009.10.042.
57. Donnelly K, Twenge JM. Masculine and feminine traits on the bem sex-role inventory, 1993–2012: A cross-temporal meta-analysis. *Sex Roles*. 2017;76(9):556-565.
58. McLean CP, Asnaani A, Litz BT, Hofmann SG. Gender differences in anxiety disorders: Prevalence, course of illness, comorbidity and burden of illness. *Journal of psychiatric research*. 2011;45(8):1027-1035.
59. Bekhuis E, Schoevers RA, van Borkulo CD, Rosmalen JG, Boschloo L. The network structure of major depressive disorder, generalized anxiety disorder and somatic symptomatology. *Psychological medicine*. 2016;46(14):2989-2998. doi: 10.1017/S0033291716001550.
60. Löwe B, Spitzer RL, Williams JB, Mussell M, Schellberg D, Kroenke K. Depression, anxiety and somatization in primary care: Syndrome overlap and functional impairment. *General hospital psychiatry*. 2008;30(3):191-199. doi: 10.1016/j.genhosppsych.2008.01.001.
61. Niles AN, O'Donovan A. Comparing anxiety and depression to obesity and smoking as predictors of major medical illnesses and somatic symptoms. *Health psychology*. 2019;38(2):172-181. doi: 10.1037/hea0000707.
62. Tak LM, Kingma EM, van Ockenburg SL, Ormel J, Rosmalen JGM. Age- and sex-specific associations between adverse life events and functional bodily symptoms in the general population. *Journal of Psychosomatic Research*. 2015;79(2):112-116. doi: 10.1016/j.jpsychores.2015.05.013.
63. Bonvanie IJ, Janssens K, Rosmalen J, Oldehinkel AJ. Life events and functional somatic symptoms: A population study in older adolescents. *Br J Psychol*. 2017;108(2):318-333. doi: 10.1111/bjop.12198.

Appendices

Appendix A: Demographic characteristics of the study population at baseline (N=150,494).

		Men (N=62,364; 41.4%)	Women (N=88,130; 58.6%)
Mean age in years (SD)		45.1 (13.0)	44.0 (12.9)
Education	Low	19,432 (31.2%)	27,869 (31.6%)
	Medium	23,562 (37.8%)	35,166 (39.9%)
	High	19,370 (31.1%)	25,095 (28.5%)
Median gender index (IQR)		0.06 (0.01-0.24)	0.96 (0.83-0.99)
Median SCL-90 SOM sumscore		14.0 (13.0-17.0)	14.0 (13.0-17.0)
Presence of chronic physical disease	No	38,404 (61.6%)	45,000 (51.1%)
	Yes	23,960 (38.4%)	43,130 (48.9%)
Employment	No	10,244 (16.6%)	19,774 (22.7%)
	Yes	51,343 (83.4%)	67,223 (77.3%)
Experienced a longterm difficulty	No	15,047 (25.2%)	17,369 (20.6%)
	Yes	44,736 (74.8%)	66,928 (79.4%)
Experienced a negative life event	No	25,101 (42.0%)	31,808 (37.7%)
	Yes	34,693 (58.0%)	52,513 (62.3%)

Appendix B: The SCL-90 SOM subscale

The Symptom CheckList-90 somatization subscale (SCL-90 SOM) is part of the Symptom CheckList-90 (SCL-90). It asks participants to score how much they were bothered or distressed by twelve somatic symptoms in the past week. The total score may range from 12–60, whereas the mean score of the SCL-90 SOM may range from 1-5.

How much in the past week were you bothered by:		**Not at all**	**A little bit**	**Moderately**	**Quite a bit**	**Extremely**
1	Headache	1	2	3	4	5
2	Dizziness	1	2	3	4	5
3	Chest pain	1	2	3	4	5
4	Lower back pain	1	2	3	4	5
5	Nausea	1	2	3	4	5
6	Painful muscles	1	2	3	4	5
7	Difficulties breathing	1	2	3	4	5
8	Feeling hot and cold alternately	1	2	3	4	5
9	Numbness or tingling in parts of your body	1	2	3	4	5
10	Feeling a lump in your throat	1	2	3	4	5
11	Weakness in body parts	1	2	3	4	5
12	Heavy arms or legs	1	2	3	4	5

Appendix C: Overview of the fit indices of the fitted models

G	Polynomial degree	NPM	Random effects	Log-Like	BIC	Entropy
1	Linear	3	0	-254,288	508,611	100%
	Linear	4	1	-190,022	380,092	100%
	Linear	6	2	-189,296	378,664	100%
2	Linear	6	0	-181,449	362,969	91.6%
	Linear	7	1	-167,150	334,384	92.3%
	Linear	9	2	-167,051	334,210	92.4%
3	Linear	9	0	-161,083	322,273	87.6%
	Linear	10	1	-167,150	334,419	45.2%
	Linear	12	2	-167,051	334,246	40.7%
4	Linear	12	0	-154,669	309,482	85.6%
	Linear	13	1	-167,150	334,456	24.7%
	Linear	15	2	-167,051	334,282	24.4%
5	Linear	15	0	-149,292	298,764	86.1%
	Linear	16	1	-167,150	334,491	18.1%
	Linear	18	2	-167,051	334,317	18.4%
6	Linear	18	0	-154,669	309,553	51.8%
	Linear	19	1			
	Linear	21	2			
7	Linear	21	0			
	Linear	22	1			
	Linear	24	2			

Reported are: the number of latent classes; the model's polynomial form; the number of estimated parameters; the presence of random effects in the model with 0 being no random effect (LCGA), 1 allowing for individual variance around the class' intercept (GMM), 2 allowing for individual variance of both the intercept and slope around the class' mean (GMM); the maximum Log-Likelihood (Log-Like); the Bayesian Information Criterion (BIC) value; the model's entropy; and for models with g≥2 classes, the *a-posteriori* classification of participants in each class and;

Participants per class (%)						Mean class membership posterior probabilities
150,494 (100%)						1
150,494 (100%)						1
150,494 (100%)						1
132,523 (88.1%)	1,7971 (11.9%)					0.98/0.92
138,431 (92.0%)	12,063 (8.0%)					0.98/0.91
138,347 (91.9%)	12,147 (8.1%)					0.98/0.90
24,711 (16.4%)	5,516 (6.7%)	120,267 (79.9%)				0.86/0.92/0.96
137,477 (91.4%)	13,017 (8.6%)	0 (0%)				0.86/0.70/0
137,160 (91.1%)	0 (0%)	13,334 (8.9%)				0.63/0/0.86
1,985 (1.3%)	8,054 (5.4%)	114,618 (76.2%)	25,837 (17.2%)			0.91/0.86/0.94/0.81
14,628 (9.7%)	0 (0%)	135,866 (90.3%)	0 (0%)			0.81/0/0.38/0
14,833 (9.9%)	0 (0%)	135,661 (90.1%)	0 (0%)			0.81/0/0.36/0
113,444 (75.4%)	1,717 (1.1%)	7,168 (4.8%)	25,954 (17.3%)	2,211 (1.5%)		0.94/0.90/0.85/0.81/0.82
15,381 (10.2%)	135,113 (89.8%)	0 (0%)	0 (0%)	0 (0%)		0.78/0.27/0/0/0
135,050 (89.7%)	154,44 (10.4%)	0 (0%)	0 (0%)	0 (0%)		0.27/0.78/0/0/0
111,525 (74.1%)	8,054 (5.4%)	28,930 (19.2%)	0 (0%)	1,985 (1.3%)	0 (0%)	0.61/0.86/0.77/0/0.91/0
						Failed to converge
						Failed to converge
						Failed to converge
						Failed to converge
						Failed to converge

the mean of posterior probabilities in each latent class. To ensure that the models did not converge on local maxima of the Log-Likelihood, all models were estimated by means of a grid search. This indicates that multiple models with different randomly selected starting values are explored. Use of a grid search ensures that the estimated parameters reflect the global maximum Log-Likelihoods. The variance/covariance matrix was constrained over latent classes and non-structured.

Appendix C: Continued

G	Polynomial degree	NPM	Random effects	Log-Like	BIC	Entropy
1	Quadratic	4	0	-254,097	508,243	100%
	Quadratic	5	1	-189,526	379,111	100%
	Quadratic	10	2	-187,738	375,595	100%
2	Quadratic	8	0	-180,602	361,299	91.5%
	Quadratic	9	1	-165,580	331,267	92.3%
	Quadratic	14	2	-187,738	375,643	0%
3	Quadratic	12	0	-159,737	319,616	87.4%
	Quadratic	13	1	-165,580	331,315	50.3%
	Quadratic	18	2	-165,042	330,298	0%
4	Quadratic	16	0	-159,737	319,664	66.8%
	Quadratic	17	1	-165,580	331,362	26.5%
	Quadratic	22	2	-165,042	330,346	0%
5	Quadratic	20	0	-153,087	306,412	78.5%
	Quadratic		1			
	Quadratic		2			

Participants per class (%)					Mean class membership posterior probabilities
150,494 (100%)					1
150,494 (100%)					1
150,494 (100%)					1
132,124 (87.8%)	18,370 (12.2%)				0.98/0.92
11,749 (7.8%)	138,745 (92.2%)				0.89/0.99
95,002 (63.1%)	55,492 (36.9%)				0.50/0.50
5,504 (3.7%)	120,203 (79.9%)	24,787 (16.5%)			0.93/0.96/0.86
138,018 (92.2%)	0 (0%)	12,476 (8.3%)			0.76/0/0.87
137,838 (91.6%)	12,656 (8.4%)	0 (0%)			0.70/0.87/0
5,499 (3.7%)	119,622 (79.5%)	0 (0%)	25,373 (16.9%)		0.93/0.82/0/0.85
137,083 (91.1%)	13,411 (8.9%)	0 (0%)	0 (0%)		0.47/0.84/0/0
136,240 (90.5%)	14,254 (9.5%)	0 (0%)	0 (0%)		0.37/0.81/0/0
0 (0%)	8,121 (5.4%)	25,704 (17.1%)	114,499 (76.1%)	2,170 (1.4%)	0/0.84/0.8/0.90/0.91
					Failed to converge
					Failed to converge

Appendix D: Individual SCL-90 SOM trajectories over time, stratified by class.

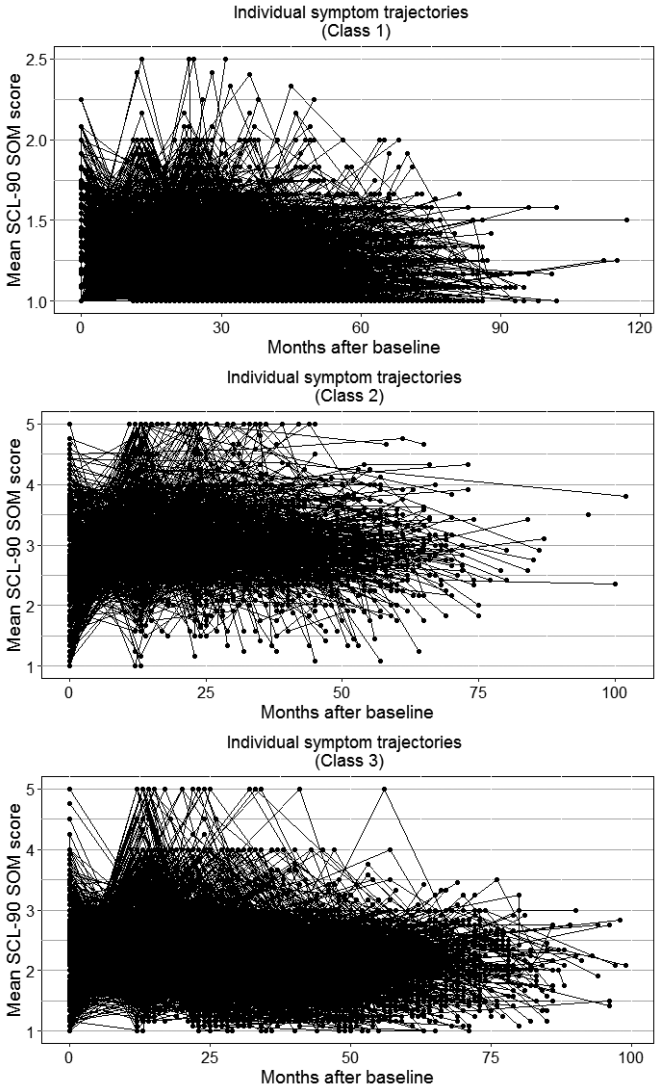


Figure D: Individuals' symptom trajectories over time, stratified by class.

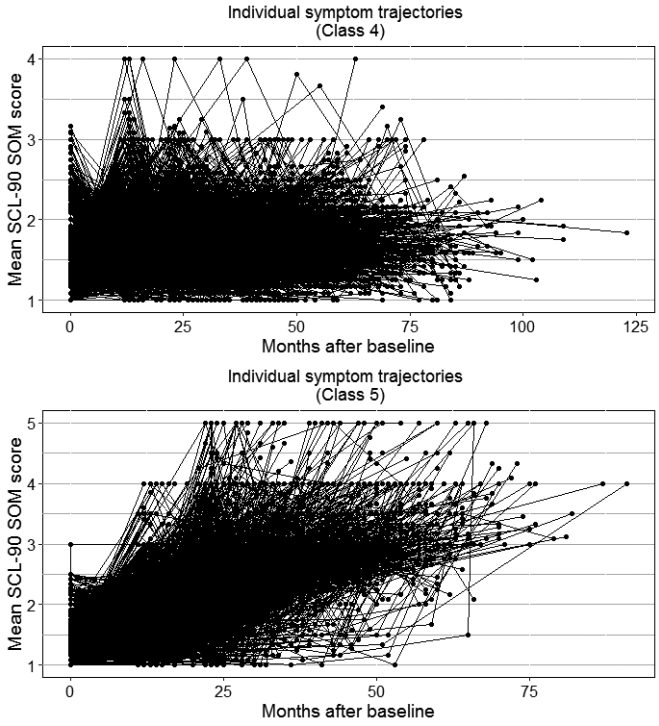


Figure D: Continued

Note: for reasons of clarity, N=5,000 participants are shown per class if the class size exceeds 5,000 participants.

Appendix E: Sensitivity Analyses

Sensitivity analyses: exploring the effect of chronic diseases on the association between femininity and high symptom severity.

		Odds ratio (95% CI)	
Predictors		**Total (N=27,183)**	
Sex	Male	1	1
	Female	1.36 (1.11-1.66)	1.31 (1.07-1.60)
Femininity		0.92 (0.73-1.16)	0.84 (0.64-1.07)
Age		1.00 (1.00-1.01)	1.00 (0.99-1.00)
Education	Low	1	1
	Medium	0.52 (0.47-0.58)	0.52 (0.47-0.58)
	High	0.28 (0.23-0.33)	0.28 (0.24-0.33)
Presence of chronic disease		n.a.	2.37 (2.10-2.67)

Sensitivity analyses: exploring the effect of chronic diseases on the association between femininity and increasing symptom severity.

		Odds ratio (95% CI)	
		Total (N=27,572)	
Sex	Male	1	1
	Female	1.30 (1.09-1.55)	1.30 (1.09-1.56)
Femininity		0.70 (0.57-0.87)	0.67 (0.54-0.84)
Age		1.02 (1.02-1.03)	1.02 (1.02-1.03)
Education	Low	1	1
	Medium	0.82 (0.74-0.91)	0.82 (0.74-0.92)
	High	0.66 (0.58-0.75)	0.65 (0.58-0.74)
Presence of chronic disease		n.a.	0.94 (0.86-1.04)

Sensitivity analyses: exploring the effect of any health-related negative life event on the association between negative life events and high symptom severity.

		Odds ratio (95% CI)	
Predictors		Total (N=27,183)	
Sex	Male	1	1
	Female	1.35 (1.10-1.67)	1.35 (1.10-1.67)
Femininity		0.84 (0.66-1.08)	0.84 (0.66-1.08)
Age		1.00 (0.99-1.00)	1.00 (0.99-1.00)
Education	Low	1	1
	Medium	0.54 (0.49-0.62)	0.55 (0.49-0.62)
	High	0.30 (0.25-0.36)	0.30 (0.25-0.36)
Presence of chronic disease		2.53 (2.22-2.88)	2.54 (2.23-2.90)
Occurrence of negative life event		2.21 (1.92-2.55)	2.12 (1.85-2.45)[a]

[a]This odds ratio reflects the association between the occurrence of non-health-related negative life and high or increasing symptom severity.

Sensitivity analyses: exploring the effect of any health-related negative life event on the association between negative life events and increasing symptom severity.

		Odds ratio (95% CI)	
		Total (N=27,572)	
Sex	Male	1	1
	Female	1.32 (1.10-1.58)	1.32 (1.10-1.58)
Femininity		0.68 (0.55-0.84)	0.68 (0.55-0.84)
Age		1.02(1.02-1.02)	1.02(1.02-1.02)
Education	Low	1	1
	Medium	0.82 (0.73-0.91)	0.82 (0.73-0.91)
	High	0.64 (0.56-0.73)	0.64 (0.56-0.73)
Presence of chronic disease		0.97 (0.88-1.06)	0.97 (0.88-1.06)
Occurrence of negative life event		1.23 (1.11-1.36)	1.23 (1.11-1.36)[a]

[a]This odds ratio reflects the association between the occurrence of non-health-related negative life and high or increasing symptom severity.

Sensitivity analyses: The associations between predictors and high, stable symptom severity over time (class 2), with no symptoms (class 1) as a reference group.

Predictors		Odds ratio (95% CI)[a]		
		Total (N=114,381)	**Men (N=51,724)**	**Women (N=62,657)**
Sex	Male	1	n.a.	n.a.
	Female	2.19 (1.67-2.86)*	n.a.	n.a.
Femininity		0.68 (0.48-0.96)*	0.71 (0.40-1.25)	0.64 (0.41-1.01)
Age		0.99 (0.98-1.00)	0.99 (0.97-1.00)	0.99 (0.98-1.00)
Education	Low	1	1	1
	Medium	0.63 (0.53-0.74)*	0.64 (0.47-0.88)*	0.63 (0.52-0.77)*
	High	0.34 (0.27-0.42)*	0.29 (0.18-0.46)*	0.36 (0.28-0.47)*
Hours of paid employment		0.99 (0.99-1.00)	1.00 (0.99-1.01)	0.99 (0.98-0.99)*
Presence of chronic disease		2.15 (1.82-2.55)*	2.57 (1.87-3.52)*	2.02 (1.66-2.47)*
Physical functioning		0.94 (0.94-0.95)*	0.94 (0.94-0.95)*	0.94 (0.94-0.94)*
Emotional wellbeing		0.95 (0.94-0.96)*	0.95 (0.94-0.96)*	0.95 (0.94-0.96)*
Self-rated health		1.06 (1.06-1.07)*	1.07 (1.05-1.08)*	1.06 (1.05-1.07)*
Negative affect		2.94 (2.47-3.52)*	2.73 (1.95-3.82)	3.03 (2.46-3.73)*
Positive affect		1.46 (1.21-1.76)*	1.23 (0.86-1.75)	1.55 (1.24-1.94)*
NEO PI R	Anger	1.48 (1.25-1.74)*	1.57 (1.16-2.12)*	1.47 (1.21-1.78)*
	Self-consciousness	0.78 (0.66-0.91)*	0.83 (0.60-1.14)	0.76 (0.63-0.91)*
	Impulsivity	0.96 (0.81-1.17)	1.19 (0.81-1.73)	0.93 (0.75-1.15)
	Vulnerability	0.77 (0.63-0.95)*	0.51 (0.34-0.76)*	0.89 (0.70-1.13)
	Self-discipline	1.20 (1.01-1.43)*	1.11 (0.79-1.56)	1.25 (1.02-1.54)*
	Competence	0.92 (0.72-1.19)	1.02 (0.63-1.67)	0.89 (0.66-1.20)
	Deliberation	0.96 (0.80-1.14)	0.91 (0.65-1.28)	0.97 (0.79-1.19)
	Excitement	0.96 (0.82-1.11)	0.88 (0.67-1.17)	0.98 (0.82-1.17)
Occurrence of negative life event		1.44 (1.19-1.74)*	2.03 (1.39-2.97)*	1.28 (1.03-1.59)*
Occurrence of long term difficulty		1.72 (1.24-2.40)*	1.42 (0.75-2.68)	1.85 (1.25-2.73)*
Presence of mood disorder		1.34 (1.12-1.61)*	1.83 (1.27-2.65)*	1.22 (0.98-1.51)
Presence of anxiety disorder		1.52 (1.23-1.88)*	1.24 (0.79-1.94)	1.64 (1.28-2.09)*

[a]Please note that the odds presented are per unit change on the scale of the predictor, thus magnitudes are not always directly comparable. *Indicates statistical significance (p<0.001). Nagelkerke's R^2 for the model including all participants, only the men and only the women allocated to class 2 and class 1 are 0.59, 0.58 and 0.59, respectively.

Sensitivity analyses: The associations between multiple predictors and increasing symptom severity (class 5) over time, with no symptoms (class 1) as a reference group.

Predictors		Odds ratio (95% CI)[a]		
		Total (N=114,770)	**Men (N=51,928)**	**Women (N=64,842)**
Sex	Male	1	n.a.	n.a.
	Female	1.76 (1.45-2.13)*	n.a.	n.a.
Femininity		0.71 (0.55-0.91)*	0.78 (0.52-1.16)	0.70 (0.50-0.98)*
Age		1.01 (1.00-1.01)	1.02 (1.00-1.02)	1.00 (0.99-1.01)
Education	Low	1	1	1
	Medium	0.74 (0.66-0.83)*	0.60 (0.49-0.74)*	0.82 (0.71-0.95)*
	High	0.46 (0.40-0.54)*	0.39 (0.30-0.51)*	0.51 (0.42-0.62)*
Hours of paid employment		1.00 (0.99-1.00)	1.00 (0.99-1.01)	0.99 (0.99-1.00)
Presence of chronic disease		1.48 (1.33-1.65)*	1.84 (1.53-2.22)*	1.34 (1.17-1.52)*
Physical functioning		0.97 (0.97-0.97)*	0.98 (0.97-0.98)	0.97 (0.97-0.97)*
Emotional wellbeing		0.99 (0.98-0.99)*	0.99 (0.98-0.99)*	0.99 (0.98-0.99)*
Self-rated health		1.02 (1.02-1.03)*	1.02 (1.02-1.03)*	1.02 (1.02-1.03)*
Negative affect		1.67 (1.47-1.90)*	1.49 (1.18-1.89)*	1.75 (1.50-2.05)*
Positive affect		1.07 (0.93-1.24)	0.81 (0.63-1.04)	1.25 (1.05-1.48)*
NEO PI R	Anger	1.32 (1.18-1.49)*	1.47 (1.20-1.80)*	1.24 (1.07-1.43)*
	Self-consciousness	0.99 (0.88-1.11)	0.94 (0.76-1.17)	1.00 (0.87-1.15)
	Impulsivity	1.25 (1.10-1.43)*	1.49 (1.17-1.90)*	1.17 (0.99-1.36)
	Vulnerability	0.96 (0.83-1.12)	0.86 (0.65-1.14)	0.97 (0.81-1.16)
	Self-discipline	1.21 (1.06-1.37)*	1.30 (1.03-1.63)*	1.15 (0.98-1.34)
	Competence	0.97 (0.80-1.17)	0.98 (0.70-1.36)	0.96 (0.76-1.21)
	Deliberation	1.08 (0.95-1.22)	0.93 (0.74-1.16)	1.15 (0.99-1.34)
	Excitement	0.96 (0.86-1.06)	1.02 (0.84-1.22)	0.94 (0.83-1.07)
Occurrence of negative life event		1.55 (1.38-1.74)*	1.34 (1.10-1.64)*	1.67 (1.44-1.93)*
Occurrence of long term difficulty		1.42 (1.20-1.66)*	1.47 (1.11-1.94)*	1.39 (1.14-1.69)*
Presence of mood disorder		1.44 (1.24-1.66)*	1.52 (1.15-2.02)*	1.41 (1.19-1.67)*
Presence of anxiety disorder		1.28 (1.07-1.54)*	1.72 (1.22-2.41)*	1.16 (0.93-1.44)

[a]Please note that the odds presented are per unit change on the scale of the predictor, thus magnitudes are not always directly comparable. *Indicates statistical significance (p<0.001). Nagelkerke's R^2 for the model including all participants, only the men and only the women allocated to class 5 and class 4 are 0.13, 0.14 and 0.12, respectively.

CHAPTER 6

Ballering, A.V., van Zon, S.K.R., Olde Hartman, T.C., & Rosmalen, J.G.M. for the Lifelines Corona Research Initiative (2022). Persistence of somatic symptoms after COVID-19 in the Netherlands: an observational cohort study. *The Lancet*, *400*(10350), 452-461.

Abstract

Background: Patients often report various symptoms after recovery from acute COVID-19. Previous studies on post- COVID-19 condition have not corrected for the prevalence and severity of these common symptoms before COVID-19 and in populations without SARS-CoV-2 infection. We aimed to analyze the nature, prevalence, and severity of long- term symptoms related to COVID-19, while correcting for symptoms present before SARS-CoV-2 infection and controlling for the symptom dynamics in the population without infection.

Methods: This study is based on data collected within Lifelines, a multidisciplinary, prospective, population-based, observational cohort study examining the health and health-related behaviors of people living in the north of the Netherlands. All Lifelines participants aged 18 years or older received invitations to digital COVID-19 questionnaires. Longitudinal dynamics of 23 somatic symptoms surrounding COVID-19 diagnoses (due to SARS-CoV-2 alpha [B.1.1.7] variant or previous variants) were assessed using 24 repeated measurements between March 31, 2020, and August 2, 2021. Participants with COVID-19 (a positive SARS-CoV-2 test or a physician's diagnosis of COVID-19) were matched by age, sex, and time to COVID-19-negative controls. We recorded symptom severity before and after COVID-19 in participants with COVID-19 and compared that with matched controls.

Findings: 76 422 participants (mean age 53.7 years [SD 12.9], 46,329 [60.8%] were female) completed a total of 883,973 questionnaires. Of these, 4231 (5.5%) participants had COVID-19 and were matched to 8462 controls. Persistent symptoms in COVID-19-positive participants at 90–150 days after COVID-19 compared with before COVID-19 and compared with matched controls included chest pain, difficulties with breathing, pain when breathing, painful muscles, ageusia or anosmia, tingling extremities, lump in throat, feeling hot and cold alternately, heavy arms or legs, and general tiredness. In 12.7% of patients, these symptoms could be attributed to COVID-19, as 381 (21.4%) of 1782 COVID-19-positive participants versus 361 (8.7%) of 4130 COVID-19-negative controls had at least one of these core symptoms substantially increased to at least moderate severity at 90–150 days after COVID-19 diagnosis or matched timepoint.

Interpretation: To our knowledge, this is the first study to report the nature and prevalence of post-COVID-19 condition, while correcting for individual symptoms present before COVID-19 and the symptom dynamics in the population without SARS-CoV-2 infection during the pandemic. Further research that distinguishes potential mechanisms driving post-COVID-19-related symptomatology is required.

Research in context

Evidence before this study: We searched PubMed, Google Scholar, and preprint repositories from November, 2019, to February, 2022, for studies published in Dutch or English that investigated the course of post-COVID-19 condition (ie, long COVID) over time, the symptoms associated with post-COVID-19 condition, and the prevalence of post-COVID-19 condition. Furthermore, we searched for studies and policy documents from (global) public health institutes (eg, WHO) that aimed to clinically define post-COVID-19 condition. A formal systematic review was not conducted. Most previous research that assessed the prevalence and symptoms associated with post-COVID-19 condition did not include an adequate control group, and so no adjustments for the prevalence of somatic symptoms in the population without COVID-19 could be made. Additionally, we found no studies that included patients' symptom prevalence before COVID-19 diagnosis; therefore, the previous studies were unable to assess whether somatic symptoms reported after a COVID-19 diagnosis were already present before SARS-CoV-2 infection. Most research was conducted in a clinical setting, disregarding post-COVID-19 condition in the general population. In the context of these shortcomings, a systematic review estimated that the median proportion of patients with at least one somatic symptom after COVID-19 was 72·5%.

Added value of this study: To our knowledge, this study is the first to include a control group matched for age, sex, and time, enabling us to adjust for symptom presence in the general population and changes herein due to public health measures and seasonal influences. Additionally, the repeated-measures nature of this study enabled us to assess symptom severity in patients with COVID-19 before they had SARS-CoV-2 infection. Therefore, we could assess whether symptom severity was truly increased after a COVID-19 diagnosis, or whether symptoms were a continuation of pre-existing symptoms. Our approach allowed for identification of core symptoms that define post-COVID-19 condition, as these are increased in severity 90–150 days after a COVID-19 diagnosis compared with patient's pre-existing symptom severity.

Implications of all the available evidence: Our unique approach allows us to present the core symptoms, namely chest pain, difficulties with breathing, pain when breathing, painful muscles, ageusia or anosmia, tingling extremities, lump in throat, feeling hot and cold alternately, heavy arms or legs, and general tiredness, which could define post-COVID-19 condition. Additionally, we offer an improved working definition of post-COVID-19 condition and provide a reliable prevalence estimate in the general population corrected for pre-existing symptoms, and symptoms in COVID-19- negative controls. Taking into account the symptoms that increased in severity and could be

attributed to COVID-19, while correcting for seasonal fluctuations and non-infectious health aspects of the pandemic on symptom dynamics, we estimated that 12·7% of patients with COVID-19 in the general population will experience persistent somatic symptoms after COVID-19. Additionally, these core symptoms have major implications for future research, as these symptoms have the highest discriminative ability to distinguish between post- COVID-19 condition and non-COVID-19-related symptoms.

Introduction

After recovery from acute COVID-19, a substantial proportion of patients continue to experience symptoms of a physical, psychological, or cognitive nature.[1] These long-term sequelae of COVID-19 have been described as the next public health disaster in the making, and there is an urgent need for empirical data informing on the scale and scope of the problem to support the development of an adequate health-care response.[2,3]

Research has been hampered by an absence of a consensus on the prevalence and nature of the post- COVID-19 condition.[2] A systematic review examining the frequency and variety of persistent symptoms after COVID-19 reported that the median proportion of patients with at least one persistent symptom was 72.5%.[4] However, this estimated prevalence largely depends on the timeframe, population, and symptoms used to define post-COVID-19 condition. The timeframe used varies from 4 weeks to more than 6 months after a COVID-19 diagnosis, with 3 months being the most commonly used.[5] Furthermore, most studies have relied on follow-up of hospitalized patients with COVID-19.[4] The vast majority of people with COVID-19, however, have mild disease and are not hospitalized,[6] and hospitalization itself is associated with somatic symptoms.[7]

Another complicating factor is that there is no consensus on the nature of the symptoms that can be attributed to COVID-19. Selection of the symptoms is crucial for charting the scale and scope of post- COVID-19 condition. However, frequently reported post-COVID-19 symptoms are also common in the general population.[4,8,9] Symptoms such as fatigue and headaches might be worsened during the pandemic also in people without COVID-19, for example, due to anxiety-induced stress or the combination of work and homeschooling.[10,11] An additional complication is that some of the symptoms reported after COVID-19 might already have been present before COVID-19 and might even reflect a pre-existing susceptibility to COVID-19 itself, rather than being a consequence of SARS-CoV-2 infection.

Therefore, detailed information about symptom dynamics before and after SARS-CoV-2 infection in the general population is needed to provide insight into the scale and scope of post-COVID-19 condition. However, such data—requiring repetitive measurements of symptom scores before and after SARS-CoV-2 infection— have not yet been reported. Furthermore, symptom dynamics need to be compared between people affected by COVID-19 and a matched sample of people without infection to be able to separate the effects of the SARS-CoV-2 infection from the effects of the pandemic, associated social restrictions, and public health measures on symptom dynamics in the general population.[12]

We aimed to analyze the nature, prevalence, and severity of long-term symptoms related to COVID-19, while correcting for symptoms present before SARS-CoV-2 infection and controlling for the symptom dynamics in the population without infection.

Methods

Study design and participants

This study is based on data collected within the Lifelines COVID-19 cohort study, an add-on study to the multidisciplinary, prospective, population-based, observational Dutch Lifelines cohort study examining the health and health-related behaviors of 167 729 people in the north of the Netherlands (>98% White, 58% female).[13,14] There were no specific inclusion criteria for the Lifelines study. Exclusion criteria for Lifelines were severe mental illness, short life expectancy (<5 years) at time of inclusion, insufficient knowledge of the Dutch language to complete questionnaires, and not being able to visit a general practitioner.[13] All Lifelines participants aged 18 years or older with a known email address received invitations to complete digital COVID-19 questionnaires.[15] We included data from 24 consecutive measurements collected in the Lifelines COVID-19 cohort study between March 31, 2020, and August 2, 2021, for which response rates varied between 28% and 49%. Initially, questionnaires were sent out weekly but from June, 2020, data were collected every 2 weeks, and from August, 2020, data were collected on a monthly basis.[15] The Lifelines cohort study and its add-on studies were approved by the Medical Ethical Committee of University Medical Center Groningen (2007/152) and participants provided written informed consent to take part. Further details on the cohort, design considerations, and recruitment procedures, as well as additional information on the COVID-19 pandemic in the Netherlands, have been published previously.[13-15]

Procedures

Participants completed digital questionnaires on multiple topics, including sociodemographics and physical and mental health during the COVID-19 pandemic. In January, 2021, the Dutch national immunization program for COVID-19 was initiated. On March 1, 2021, only 3.7% of the total Lifelines study population was fully vaccinated, increasing to 9.8% by the end of April, 2021, when the last COVID-19 cases for the current study were included. Until July, 2021, the alpha (B.1.1.7) SARS-CoV-2 variant was dominant in the Netherlands. Participants' COVID-19 positivity was defined as either a positive SARS-CoV-2 test or a physician's diagnosis of COVID-19, which was based on the evolving clinical case definition issued by the Dutch Institute for Public Health and the Environment. Physician diagnosis of COVID-19 was included as positivity because SARS-CoV-2 testing in the Netherlands was strongly restricted up until August, 2020.[6] We only included participants' first SARS-CoV-2 infections.

We analyzed 23 symptoms: headache, dizziness, chest pain, back pain, nausea, painful muscles, difficulties with breathing, feeling hot and cold alternately, tingling extremities, lump in the throat, general tiredness, heavy arms or legs, pain when breathing, runny nose, sore throat, dry cough, wet cough, fever, diarrhea, stomach pain, ageusia or anosmia, sneezing, and itchy eyes. The first 12 of these symptoms were derived from the validated Symptom CheckList-90 Somatization (SCL-90 SOM) subscale,[16] which has been shown to have sufficient measurement invariance, making it suitable to assess symptoms repeatedly over time.[17] The remainder of the symptoms were added as these were considered to be related to COVID-19 at the start of the study. All symptoms were assessed using an ordinal 5-point Likert scale that answered to what extent participants were bothered by the respective symptom (1=not at all, 5=extremely) in the past 7 days. The timeframe was changed to the past 14 days when questionnaires were sent out every 2 weeks and monthly. The item assessing sneezing was introduced in the second questionnaire. Stomach pain and diarrhea were assessed by a combined item in the first two questionnaires, but thereafter these were assessed by separate items. We included these first two measurements for both stomach pain and diarrhea. Presence of symptoms was defined by a score of at least 3 (ie, moderately bothered by the symptom).

To assess persistence of somatic symptoms after a COVID-19 diagnosis, we first calculated participants' individual mean pre-COVID-19 score per symptom. We excluded reports on symptom severity collected in the week before COVID-19 diagnosis, as increased symptom severity might have prompted participants to seek SARS-CoV-2 testing or a physician's diagnosis. If no information on symptom severity before the COVID-19 diagnosis was available, participants were excluded from the analyses.

Statistical analysis

Characteristics of the study population, including age, level of education, and presence of chronic disease (***Appendix A***), are provided as absolute numbers with concomitant percentages. If appropriate, information on continuous measurements are provided as means with SD. Data were examined for normality using Q-Q plots and histograms. All information, except for age and sex, was self-reported by participants.

We describe COVID-19-positive participants' moving average symptom report, including the moving average's SE, over time, stratified by symptom and participants' sex. The moving average was based on an interval of 28 days to avoid fluctuations in symptoms resulting from differences in subsamples of the cohort completing the questionnaire on a specific day. We randomly matched COVID-19-positive participants with COVID-19-negative controls (1:2), by sex (male or female), age (split at the median, ≤52 years or ≥53 years), and time of completing questionnaires that indicated a COVID-19 diagnosis (measurement wave, range: 1–24). We matched for time to account for the variation in symptom burden during the pandemic in the population without COVID-19 due to seasonal effects and non-infectious pandemic consequences.

We describe the relative frequency of COVID-19-positive participants and COVID-19-negative matched controls in whom mean symptom severity from 90 to 150 days after COVID-19 was at least moderate (ie, score of ≥3), stratified per symptom. Additionally, we describe the relative frequency of participants in whom mean symptom severity from 90 to 150 days after a COVID-19 diagnosis was increased by at least 1 point compared with before COVID-19, resulting in at least moderate symptom severity (ie, score of ≥3), stratified per symptom. We further assessed whether the presence of symptoms that increased substantially to at least moderate symptom severity from 90 to 150 days after COVID-19 or matched time differed in distribution between COVID-19-positive participants and COVID-19-negative controls via χ^2 tests. We maintained a two-sided α level of $p<0.001$ (0.05 divided by 50; 23 symptoms and two [sub]totals, each tested two times) to correct for the number of performed tests. We assessed mean symptoms at 90–150 days after COVID-19, based on at least one questionnaire, in concordance with the recently proposed WHO clinical case definition for post-COVID-19 condition that states symptoms occur usually 3 months from the onset of COVID-19 and last for at least 2 months. The 1-point difference was assessed as this is the minimal change participants could indicate on the 5-point Likert scale.

Furthermore, in sensitivity analyses, we assessed the prevalence of symptoms at 3 months after COVID-19 restricted to people who had a positive SARS-CoV-2 test. IBM SPSS (version 25) was used to perform all analyses. In compliance with the SAGER guidelines, we report our findings stratified by participants' sex if appropriate.[18]

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

76 422 participants (mean age 53.7 years [SD 12.9], 46 329 [60.8%] were female) completed a total of 883 973 questionnaires. Of these, 4231 (5.5%) participants were COVID-19 positive (mean age 52.4 years [SD 11.7], 2779 [65.7%] were female; ***Table 1***; ***Appendix B***); they completed 62 224 questionnaires (***Appendix C***).

Table 1. Characteristics of the COVID-19-positive study population

		Male participants (n=1,452)	Female participants (n=2,779)
Age, years		54.3 (11.5%)	51.4 (11.7%)
BMI, kg/m²		26.6 (3.7%)	26.3 (4.9%)
Educational level	Low	193 (13.3%)	274 (9.9%)
	Medium	694 (47.8%)	1484 (53.4%)
	High	532 (36.6%)	896 (32.2%)
	Missing	33 (2.3%)	125 (4.5%)
Chronic disease[a]	Absent	1225 (84.4%)	2209 (79.5%)
	Present	110 (7.6%)	287 (10.3%)
	Missing	117 (8.1%)	283 (10.2%)
Smoking	No	1,356 (93.4%)	2543 (91.5%)
	Yes	63 (4.3%)	143 (5.1%)
	Missing	33 (2.3%)	93 (3.3%)
Method of diagnosis	Physician's diagnosis	297 (20.5%)	602 (21.7%)
	Positive SARS-CoV-2 test	1155 (79.5%)	2177 (78.3%)
Hospitalization		72 (5.0%)	70 (2.5%)

Data are mean (SD) or n (%). [a]See the appendix for the full list of included chronic diseases. The characteristics of the COVID-19-negative controls compared with COVID-19-positive participants are provided in the appendix.

Female COVID-19-positive participants completed a median of 17 questionnaires (IQR 8–23), male COVID-19-positive participants completed a median of 18 (IQR 9–23). The maximum follow-up time was 484 days after COVID-19 diagnosis (median 101 days [IQR 43–199]). COVID-19-positive participants were matched to 8462 COVID-19-negative controls who together completed 140 810 questionnaires (***Appendix C***). Both male and female control participants completed a median of 20 questionnaires (IQR 12–24) each. The maximum follow-up time of control participants was 481 days after their matched timepoint (median 104 days [IQR 46–201]). The sex-stratified 28-day moving average of control participants' mean sum score of all 23 assessed symptoms is shown in the ***Appendix D***. Men were more frequently hospitalized due to COVID-19 than women (5.0% of male vs 2.5% of female COVID-19- positive participants).

Visual inspection of symptom dynamics over time indicated that almost all assessed symptoms showed an increase in severity in COVID-19-positive participants compared with controls during the acute phase of COVID-19 (***Figures 1–3***). Diarrhea and stomach pain, as well as cold-like symptoms including sneezing, wet and dry cough, runny nose, fever, and sore throat on average returned to pre-COVID-19 severity within 50 days of a COVID-19 diagnosis, which suggests that these symptoms were predominantly present during the acute phase of the disease (***Figure 1***).

Symptoms that were more severe in COVID-19-positive participants 90–150 days after COVID-19 compared with symptom scores before COVID-19 and compared with matched controls (ie, the core symptoms of post- COVID-19 condition) included: cardiopulmonary symptoms (chest pain, difficulties with breathing, and pain when breathing), musculoskeletal symptoms (painful muscles), sensory symptoms (ageusia or anosmia, tingling extremities, lump in throat, and feeling hot and cold alternately), and general symptoms (heavy arms or legs, and general tiredness; ***Figure 2***).

These symptoms differed based on both visual inspection of symptom dynamics and on the significance of the difference in distribution of symptoms that increased substantially to at least moderate severity in COVID-19- positive participants and control participants (***Table 2***). Mean severity for these symptoms appeared to have reached a plateau at 3 months, with no further decline in mean severity thereafter. Symptoms that were not significantly increased in mean severity at 90–150 days after a COVID-19 diagnosis included headache, itchy eyes, dizziness, back pain, and nausea (***Figure 3***).

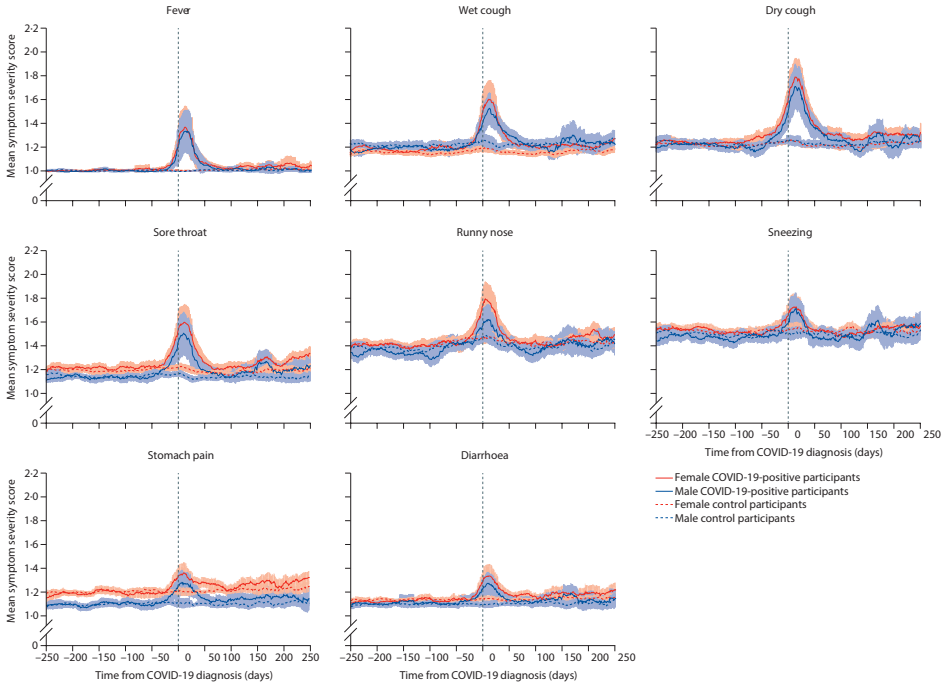


Figure 1. Acute symptoms. The shaded areas represent the SE of the moving average.

Visual inspection of the core symptoms suggests that in many of these symptoms, including lump in throat, heavy arms or legs, general tiredness, and feeling hot and cold alternately, sex differences were present. Female COVID-19-positive participants showed a longer persistence of increased symptom severity after COVID-19 than male COVID-19-positive participants (***Figure 2***). A similar pattern was observed in acute symptoms, such as dry cough, stomach pain, and diarrhea (***Figure 1***), and in all symptoms that were not significantly increased in severity at 90–150 days after a COVID-19 diagnosis, except for back pain. ***Table 2*** shows the frequencies of COVID-19-positive participants and controls that had symptoms of at least moderate severity at 90–150 days after COVID-19 or matched timepoint. In total, 790 (40.7%) of 1942 COVID-19- positive participants had at least one symptom of moderate severity at 90–150 days, compared with 1275 (29.3%) of 4353 controls. Painful muscles and back pain were the most frequent symptoms in both COVID-19-positive participants (13.5% and 10.8%, respectively) and controls (8.7% and 9.5%, respectively). This analysis, however, did not consider symptom severity before COVID-19.

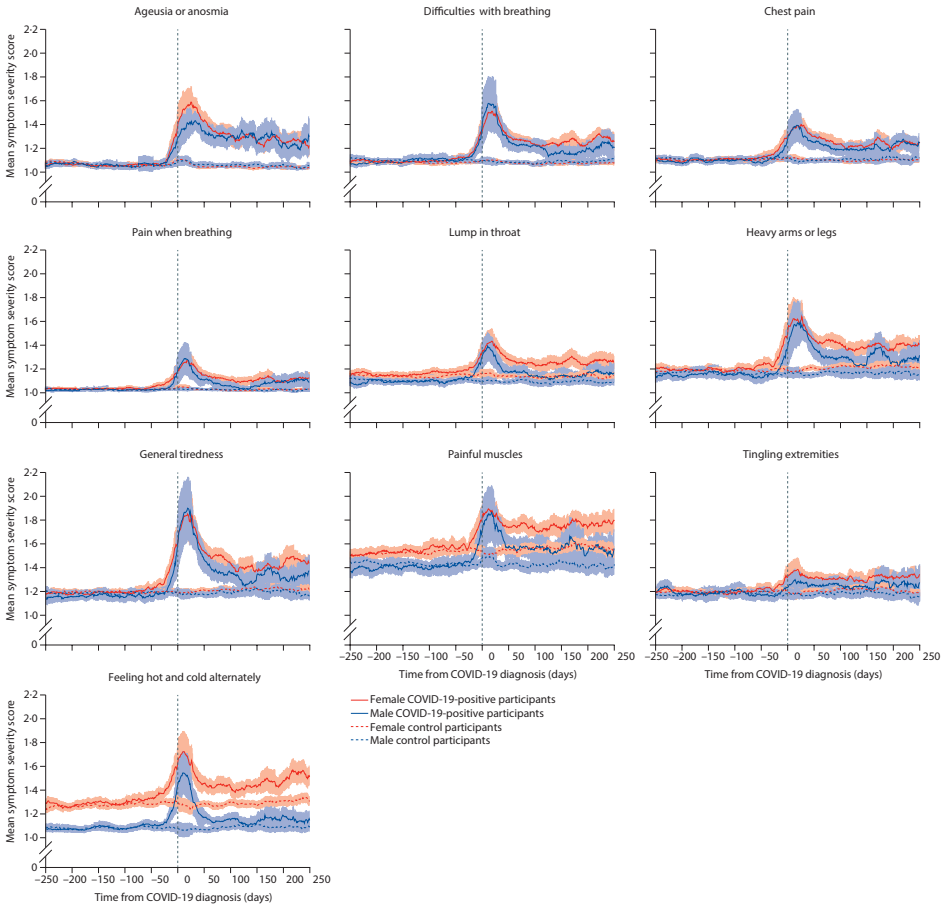


Figure 2. Core symptoms. The shaded areas represent the SE of the moving average.

A greater proportion of COVID-19-positive participants had a substantial increase in symptom severity resulting in moderate symptom severity of at least one symptom at 90–150 days after COVID-19 diagnosis than control participants during the same period (526 [29.6%] of 1782 participants vs 749 [18.1%] of 4130; ***Table 2***). Ageusia or anosmia (135 [7.6%] of 1782 participants), painful muscles (130 [7.3%]) and general tiredness (88 [4.9%]) were most frequently increased to moderate severity in COVID-19-positive participants, while they were increased in 17 (0.4%), 134 (3.2%), and 87 (2.1%) control participants, respectively. The prevalence of ageusia or anosmia of increased severity (7.6%) was 19 times greater in COVID-19-positive participants than in controls (0.4%). Sensitivity analyses in which participants with a physician's diagnosis of COVID-19 were excluded (including only those with a positive SARS-CoV-2 test) showed similar results (***Appendix E***).

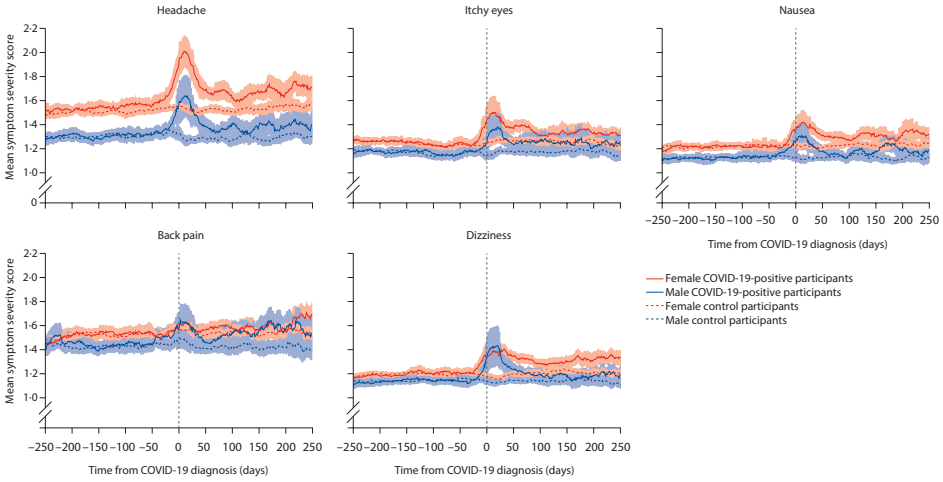


Figure 3. Other symptoms. The shaded areas represent the SE of the moving average.

Restricting the definition of post-COVID-19 condition to core symptoms (***Figure 2***) showed that 381 (21.4%) of 1782 COVID-19-positive participants versus 361 (8.7%) of 4130 controls had at least one symptom substantially increased to at least moderate severity (χ^2 [df=1] 181.1; p<0.0001; denominators based on participants with data available for at least 7 days before their SARS-CoV-2 infection or matched timepoint and 90–150 days after their COVID-19 diagnosis or matched timepoint). This finding implies that in 12.7% of patients with COVID-19, the increased core symptoms with moderate severity at 3 months after COVID-19 could be attributed to SARS-CoV-2 infection. Including all assessed symptoms in the definition decreased the prevalence of participants with an increase in symptom severity only slightly (to 11.5%), but resulted in a loss of sensitivity for symptoms that can be attributed to SARS-CoV-2 (ie, the ratio between patients with symptoms due to SARS-CoV-2 infection and those with unrelated symptoms was 2.5 for the core set of symptoms vs 1.6 when including all symptoms).

Table 2. Frequencies of participants who had presence of, or a substantial increase to, symptoms of at least moderate severity at 90-150 days after COVID-19 diagnosis or matched timepoint.

Symptom	Presence of symptom of at least moderate severity		Substantial increase of symptom severity to at least moderate severity	
	Controls (n=4353)	COVID-19 (n=1942)	Controls (n=4130)	COVID-19 (n=1782)
Ageusia/anosmia	37 (0.8%)	158 (8.1%)*	17 (0.4%)	135 (7.6%)*
Difficulties when breathing	38 (0.9%)	68 (3.5%)*	21 (0.5%)	43 (2.4%)*
Chest pain	44 (1.0%)	63 (3.2%)*	24 (0.6%)	43 (2.4%)*
Pain when breathing	13 (0.3%)	20 (1.0%)*	<10 (<0.2%)	16 (0.9%)*
Lump in throat	59 (1.4%)	61 (3.1%)*	24 (0.6%)	42 (2.4%)*
Heavy arms/legs	130 (3.0%)	126 (6.5%)*	65 (1.6%)	75 (4.2%)*
General tiredness	159 (3.7%)	136 (7.0%)*	87 (2.1%)	88 (4.9%)*
Painful muscles	378 (8.7%)	262 (13.5%)*	134 (3.2%)	130 (7.3%)*
Tingling extremities	145 (3.3%)	98 (5.0%)*	65 (1.6%)	52 (2.9%)*
Fever	19 (0.4%)	16 (0.8%)	18 (0.4%)	12 (0.7%)
Wet cough	83 (1.9%)	58 (3.0%)	40 (1.0%)	28 (1.6%)
Dry cough	81 (1.9%)	50 (2.6%)	43 (1.0%)	28 (1.6%)
Headache	239 (5.5%)	166 (8.5%)*	111 (2.7%)	76 (4.3%)
Itchy eyes	143 (3.3%)	96 (4.9%)*	78 (1.9%)	51 (2.9%)
Alternately feeling hot/cold	155 (3.6%)	112 (5.8%)*	70 (1.7%)	63 (2.5%)*
Sore throat	84 (1.9%)	48 (2.5%)	51 (1.2%)	29 (1.6%)
Runny nose	217 (5.0%)	110 (5.7%)	94 (2.3%)	50 (2.8%)
Nausea	128 (2.9%)	72 (3.7%)	74 (1.8%)	37 (2.1%)
Sneezing	210 (4.8%)	101 (5.2%)	74 (1.9%)†	35 (2.1%)‡
Back pain	413 (9.5%)	210 (10.8%)	182 (4.4%)	88 (4.9%)
Stomach pain	108 (2.5%)	53 (2.7%)	58 (1.4%)	25 (1.4%)
Dizziness	93 (2.1%)	46 (2.4%)	56 (1.4%)	25 (1.4%)
Diarrhea	80 (1.8%)	38 (2.0%)	52 (1.3%)	19 (1.1%)
Total	1,275 (29.3%)	790 (40.7%)*	749 (18.1%)	526 (29.6%)*

Data are n (%). Symptoms are ordered according to their relative increase in frequency in OCVID-19-positive participants compared with controls. A substantial increase in severity was defined as an increase in symptom severity of at least 1 point on the 5-point scale. *p<0.001. †n=3988; sneezing was assessed in 23 surveys instead of 24. ‡n=1704; sneezing was assessed in 23 surveys instead of 24.

Discussion

This study shows post-COVID-19 condition might occur in about one out of eight people with COVID-19 in the general population. Core symptoms of post-COVID-19 condition include chest pain, difficulties with breathing, lump in throat, pain when breathing, painful muscles, heavy arms or legs, ageusia or anosmia, feeling hot and cold alternately, tingling extremities, and general tiredness. To our knowledge, this is the first study to provide a reliable assessment of the prevalence of post-COVID-19 condition, while correcting for individual symptoms present before SARS-CoV-2 infection and for the dynamics of symptoms reported by sex-matched and age-matched controls without infection in the same period during the pandemic. This corrected prevalence remained nearly unaltered irrespective of the use of the core symptoms versus a broader range of symptoms as a definition of post-COVID-19 condition. However, when including a broader range of symptoms, the ratio between patients with symptoms due to SARS-CoV-2 infection and those with unrelated symptoms decreased. Increased knowledge on both the nature of the core symptoms and the prevalence of post-COVID-19 condition in the general population represents a major step forward in our ability to design studies that ultimately inform an adequate health-care response to the long-term sequelae of COVID-19.

The major strengths of this study are the large sample size of COVID-19-positive participants identified in a general population cohort, as well as the multiple repeated measurements of symptom severity in the participants. This allowed for the calculation of pre-COVID-19 symptom severity in each participant. In addition, we were able to compare COVID-19-positive participants' symptom severity with controls matched by sex and age who provided measurements at the same time period as the cases. Finally, the SCL-90 SOM subscale is a validated instrument, suitable for assessing symptoms in large-scale cohort studies. The addition of other COVID-19-related symptoms allowed for detailed insights into participants' symptom dynamics.

Before interpreting the results, some limitations of this study should be acknowledged. First, COVID-19 cases can be asymptomatic and remain undetected.[8] Therefore, the prevalence of COVID-19 in this study might have been underestimated. Second, the assessed symptoms were included in the Lifelines COVID-19 cohort study at the beginning of the pandemic. Although at that time these symptoms were considered to be related to COVID-19, other symptoms such as cognitive symptoms (eg, brain fog) and post-exertional malaise were identified later during the pandemic as potentially relevant for a working definition of post-COVID-19 condition.[7] Third, as all participants in the Lifelines COVID-19 cohort study were aged 18 years or older, we could not assess

pediatric post- COVID-19 condition. Fourth, the exact date of COVID-19 diagnosis was unknown; we therefore used the date of the first questionnaire in which COVID-19 positivity was indicated as date of diagnosis. This might have led to an underestimation of post-COVID-19 time. Lastly, as this study was conducted in the northern region of the Netherlands, these results might not be generalizable to other areas.

Multiple studies have assessed the persistence of somatic symptoms after COVID-19, with timeframes of follow-up varying from 21 days to 6 months.[4,19] Some studies included participants from post-COVID-19 support groups or predominantly patients who were hospitalized, leading to biased results.[20,21] A systematic review analyzed 11 studies that assessed the persistence of symptoms 90–180 days after COVID-19 in outpatients.[19] The sample sizes ranged from 59 to 2915 patients with COVID-19 and the number of assessed symptoms ranged from six to 21. The most prevalent symptom was fatigue (11–42% of patients), followed by dyspnea (8–37%), painful muscles (7–24%), and ageusia or anosmia (3–24%). Thoracic pain was reported in 3–14% of patients at 90–180 days after COVID-19. Although we found similar prevalence rates for some of these symptoms, we also showed that these rates were lower when patients' symptom severity before COVID-19 was taken into account. Additionally, we showed that the most prevalent symptoms are not the most distinctive symptoms for post-COVID-19 condition. Furthermore, many studies with clinical cohorts did not include a matched control group and were therefore unable to distinguish between effects of SARS-CoV-2 infection and those of the pandemic on symptoms.[12] Studies that included a control group could not distinguish between symptoms resulting from a SARS-CoV-2 infection and pre-existing symptoms. A large study that included 106 578 patients with COVID-19 and matched controls with influenza, which assessed the persistence of seven somatic symptoms at 90–180 days after diagnosis, found that somatic symptoms, such as headache, chest pain, and fatigue, were more frequently present in patients with COVID-19 than in the controls.[22] The study found higher prevalence rates for most assessed somatic symptoms than our study—for example, breathing difficulties occurred in 7.9% of patients with COVID-19 and chest pain occurred in 5.7%. Painful muscles was the only symptom that was less frequently reported (1.5% of patients). The difference in observed prevalence rates might be explained by the previous study only including patients with COVID-19 who sought help for their persistent symptoms from a health-care provider, and not adjusting for patients' symptoms before COVID-19.

Additionally, a study in France that included 1091 SARS-CoV-2-positive participants and 25 732 controls suggested that the belief of being infected with SARS-CoV-2 was more strongly associated with the severity of symptoms 8 weeks after SARS-CoV-2 infection than laboratory confirmed COVID-19 diagnosis.[23] This conclusion is

potentially stigmatizing,[24] and the study has some limitations. First, serological assays were used to detect SARS-CoV-2 infection, but patients affected by post-COVID-19 condition might have lower antibody responses.[25] Second, the cross-sectional nature of the study with retrospective assessments is problematic, as persistent physical symptoms might have confounded recall of past illness and thus the belief in having been infected. Third, confounding by other viruses might have occurred, which might have caused both the belief of having been infected with SARS-CoV-2 and the persistent symptoms. Our study overcame these limitations by performing sensitivity analyses restricted to participants with a COVID-19 diagnosis based on a positive SARS-CoV-2 test and by the study's prospective design. Nevertheless, our study cannot provide definitive information on the underlying mechanisms driving post-COVID-19-related symptoms. Therefore, additional research assessing the causes of post-COVID-19-related symptoms is required.

To our knowledge, this is the first study that is able to identify which persistent symptoms are particularly related to SARS-CoV-2 infection, and we used these core symptoms of post-COVID-19 condition for an empirically based working definition of the condition. Notably, in the absence of adequate control data, case definitions might be biased towards highly prevalent symptoms. Experts in a WHO Delphi procedure constructed a case definition that identified fatigue and dyspnea as the most important symptoms of post-COVID-19 condition (78% of the panel agreed on their importance for the case definition).[26] Our empirical analyses showed that these were among the core symptoms, but the most distinctive symptoms also included chest pain and ageusia or anosmia (considered important for the case definition by 55% and 57% of the Delphi panel, respectively). Additionally, tingling extremities were considered important by merely 39% of the experts, while 56% considered headache to be important for the case definition. Our results, however, suggest that tingling extremities is a core symptom whereas headache is not related to SARS-CoV-2 infection. These differences clearly show the importance of longitudinal cohort studies in the general population with pre-infection data and controls without infection to study the scale and scope of post-COVID-19 condition.

Furthermore, although sex differences are known to be present in persistent somatic symptoms of COVID-19, this is the first study of our knowledge to stratify symptom dynamics by sex both before and after COVID-19. Multiple somatic symptoms -for example, feeling hot and cold alternately, lump in throat, and general tiredness- were shown to be more severe after COVID-19 in women than in men, compared with controls. Research has shown that women report more severe common somatic symptoms than men and that these symptoms are more frequently persistent.[27-29]

Multiple explanations have been proposed for this phenomenon. First, women are thought to have a heightened sensitivity to pain compared with men, due to biological differences rooted in, among others, sex hormones and genotype.[30] Second, women might be more aware of bodily sensations than men, allowing for an easier and earlier perception of somatic symptoms in women than in men.[29] However, the female preponderance in symptom experience is not only due to differences in biology (ie, sex), but also in societal expectations of women and men (ie, gender roles).[27,28] Feminine gender roles, for example, are thought to be associated with poorer access to health care, which might also explain health-related gender differences.[31]

A list of empirically validated core symptoms of post- COVID-19 condition, used for a working definition of the condition, is essential to adequately study pathophysiological mechanisms,[2] which is especially important given the risk of simple psychogenic explanations and the resulting consequences for patients.[24] Our results support a working definition at least based on the core symptoms, given the improved sensitivity ratio between cases and controls compared with a broader definition. These core symptoms were increased at 3–5 months after COVID-19, and are likely to limit functioning, prompt help-seeking, and have plausible underlying pathophysiological mechanisms. Nevertheless, research shows that COVID-19 might also affect brain functioning and mental health.[32,33] Therefore, future research should not overlook mental health symptoms (eg, depression and anxiety symptoms), nor additional post-infectious symptoms that were not assessed in this study (eg, brain fog, insomnia, and post-exertional malaise). Additionally, future intersectional research should assess how ethnicity, gender, age, socioeconomic status, other social identities, and the presence of underlying chronic diseases are associated with symptom dynamics surrounding COVID-19 and risk of post-COVID-19 condition. Further research will focus on the clustering of COVID-19 symptoms in participants, and whether symptom clusters are associated with subtypes and distinct pathophysiological mechanisms underlying post-COVID-19 condition. We will also study genetic and environmental risk factors, and how post- COVID-19 condition affects (work) functioning and wellbeing. Additionally, as research suggests that vaccination before SARS-CoV-2 infection only partly mitigates the risk of long-term symptom sequelae 6 months after COVID-19,[34] further studies should assess the effect of SARS-CoV-2 vaccination and the timing thereof, and the effect of SARS-CoV-2 variants, on symptom dynamics in both adults and children. In conclusion, we present a starting point for core symptoms that could define post-COVID-19 condition, offer an improved working definition of post-COVID-19 condition, and provide a reliable prevalence estimate in the general population of the northern region of the Netherlands corrected for pre-existing symptoms and symptoms in participants without infection. Taking into account those symptoms

that increased in severity and could be attributed to COVID-19, while correcting for seasonal fluctuations and non-infectious health aspects of the pandemic on symptom dynamics,[2,5,12] we found that about one in every eight patients are affected by persistent symptoms after COVID-19. This finding shows that post-COVID-19 condition is an urgent problem with a mounting human toll.

References

1. Callard F, Perego E. How and why patients made long covid. *Social Science & Medicine*. 2021;268:113426. doi: 10.1016/j.socscimed.2020.113426.
2. Phillips S, Williams MA. Confronting our next national health disaster — long-haul covid. *New England Journal of Medicine*. 2021;385:577-579. doi: 10.1056/NEJMp2109285.
3. Crook H, Raza S, Nowell J, Young M, Edison P. Long covid—mechanisms, risk factors, and management. *BMJ*. 2021;374:n1648.
4. Nasserie T, Hittle M, Goodman SN. Assessment of the frequency and variety of persistent symptoms among patients with COVID-19: A systematic review. *JAMA Network Open*. 2021;4(5):e2111417. doi: 10.1001/jamanetworkopen.2021.11417.
5. Yong SJ. Long COVID or post-COVID-19 syndrome: Putative pathophysiology, risk factors, and treatments. *Infectious Diseases*. 2021;53(10):737-754. doi: 10.1080/23744235.2021.1924397.
6. Ballering AV, Oertelt-Prigione S, LCRI, Olde Hartman TC, Rosmalen JGM. Sex and gender-related differences in COVID-19 diagnoses and SARS-CoV-2 testing practices during the first wave of the pandemic: The dutch lifelines COVID-19 cohort study. *Journal of Women's Health (Larchmt.)*. 2021;30(12):1686-1692. doi: 10.1089/jwh.2021.0226.
7. Al-Aly Z, Xie Y, Bowe B. High-dimensional characterization of post-acute sequelae of COVID-19. *Nature*. 2021;594(7862):259-264. doi: 10.1038/s41586-021-03553-9.
8. Fernández-de-Las-Peñas C, Palacios-Ceña D, Gómez-Mayordomo V, Cuadrado ML, Florencio LL. Defining post-COVID symptoms (post-acute COVID, long COVID, persistent post-COVID): An integrative classification. *International Journal of Environmental Research and Public Health*. 2021;18(5):2621. doi: 10.3390/ijerph18052621.
9. Acevedo-Mesa A, Tendeiro JN, Roest A, Rosmalen JGM, Monden R. Improving the measurement of functional somatic symptoms with item response theory. *Assessment*. 2020;28(8):1960-1970. doi: 10.1177/1073191120947153.
10. Bekhuis E, Schoevers RA, Van Borkulo CD, Rosmalen J, Boschloo L. The network structure of major depressive disorder, generalized anxiety disorder and somatic symptomatology. *Psychological Medicine*. 2016;46(14):2989-2998. doi: 10.1017/S0033291716001550.
11. Janssens KAM, Rosmalen JGM, Ormel J, Van Oort FVA, Oldehinkel AJ. Anxiety and depression are risk factors rather than consequences of functional somatic symptoms in a general population of adolescents: The TRAILS study. *Journal of Child Psychology and Psychiatry*. 2010;51(3):304-312. doi: 10.1111/j.1469-7610.2009.02174.x.
12. Amin-Chowdhury Z, Ladhani SN. Causation or confounding: Why controls are critical for characterizing long COVID. *Nature medicine*. 2021. doi: 10.1038/s41591-021-01402-w.
13. Scholtens S, Smidt N, Swertz MA, et al. Cohort profile: LifeLines, a three-generation cohort study and biobank. *International Journal of Epidemiology*. 2015;44(4):1172-80. doi: 10.1093/ije/dyu229.
14. Klijs B, Scholtens S, Mandemakers JJ, Snieder H, Stolk RP, Smidt N. Representativeness of the LifeLines cohort study. *PLoS ONE*. 2015;10(9):e0137203. doi: 10.1371/journal.pone.0137203.
15. McIntyre K, Lanting P, Deelen P, et al. Lifelines COVID-19 cohort: Investigating COVID-19 infection and its health and societal impacts in a dutch population-based cohort. *BMJ Open*. 2021;11:e044474. doi: 10.1136/bmjopen-2020-044474.
16. Zijlema WL, Stolk RP, Löwe B, Rief W, White PD, Rosmalen JGM. How to assess common somatic symptoms in large-scale studies: A systematic review of questionnaires. *Journal of Psychosomatic Research*. 2013;74(6):459-68. doi: 10.1016/j.jpsychores.2013.03.093.

17. Rytilä-Manninen M, Fröjd S, Haravuori H, et al. Psychometric properties of the symptom checklist-90 in adolescent psychiatric inpatients and age- and gender-matched community youth. *Child and Adolescent Psychiatry and Mental Health*. 2016;10(23):23. doi: 10.1186/s13034-016-0111-x.
18. Heidari S, Babor TF, De Castro P, Tort S, Curno M. Sex and gender equity in research: Rationale for the SAGER guidelines and recommended use. *Research Integrity and Peer Review*. 2016;1:2. doi: 10.1186/s41073-016-0007-6.
19. Nguyen NN, Hoang VT, Dao TL, Dudouet P, Eldin C, Gautret P. Clinical patterns of somatic symptoms in patients suffering from post-acute long COVID: A systematic review. *European Journal of Clinical Microbiology & Infectious Diseases*. 2022;41(4):515-545. doi: 10.1007/s10096-022-04417-4.
20. Davis HE, Assaf GS, McCorkell L, et al. Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. *EClinicalMedicine*. 2021:101019. doi: 10.1016/j.eclinm.2021.101019.
21. Goërtz YMJ, Herck MV, Delbressine JM, et al. Persistent symptoms 3 months after a SARS-CoV-2 infection: The post-COVID-19 syndrome? *ERJ open research*. 2020;6(4):542. doi: 10.1183/23120541.00542-2020.
22. Taquet M, Dercon Q, Luciano S, Geddes JR, Husain M, Harrison PJ. Incidence, co-occurrence, and evolution of long-COVID features: A 6-month retrospective cohort study of 273,618 survivors of COVID-19. *PLoS medicine*. 2021;18(9):e1003773. doi: 10.1371/journal.pmed.1003773.
23. Matta J, Wiernik E, Robineau O, et al. Association of self-reported COVID-19 infection and SARS-CoV-2 serology test results with persistent physical symptoms among french adults during the COVID-19 pandemic. *JAMA Internal Medicine*. 2022;182(1):19-25. doi: 10.1001/jamainternmed.2021.6454.
24. Ballering A, Olde Hartman T, Rosmalen J. Long COVID-19, persistent somatic symptoms and social stigmatisation. *Journal of Epidemiology and Community Health*. 2021;75(6):603. doi: 10.1136/jech-2021-216643.
25. García-Abellán J, Padilla S, Fernández-González M, et al. Antibody response to SARS-CoV-2 is associated with long-term clinical outcome in patients with COVID-19: A longitudinal study. *J Clin Immunol*. 2021;41(7):1490-1501. doi: 10.1007/s10875-021-01083-7.
26. Soriano JB, Murthy S, Marshall JC, Relan P, Diaz JV. A clinical case definition of post-COVID-19 condition by a delphi consensus. *The Lancet Infectious Diseases*. 2022;22(4):e102-e107. doi: 10.1016/S1473-3099(21)00703-9.
27. Ballering AV, Bonvanie IJ, Olde Hartman TC, Monden R, Rosmalen JGM. Gender and sex independently associate with common somatic symptoms and lifetime prevalence of chronic disease. *Social Science & Medicine*. 2020;253:112968. doi: 10.1016/j.socscimed.2020.112968.
28. Ballering AV, Wardenaar KJ, olde Hartman TC, Rosmalen JGM. Female sex and femininity independently associate with common somatic symptom trajectories. *Psychological Medicine*. 2022;52(11):2144-2154. doi: 10.1017/S0033291720004043.
29. Barsky AJ, Peekna HM, Borus JF. Somatic symptom reporting in women and men. *Journal of General Internal Medicine*. 2001;16(4):266-275. doi: 10.1046/j.1525-1497.2001.00229.x.
30. Bartley EJ, Fillingim RB. Sex differences in pain: A brief review of clinical and experimental findings. *Br J Anaesth*. 2013;111(1):52-58. doi: 10.1093/bja/aet127.
31. Pelletier R, Humphries KH, Shimony A, et al. Sex-related differences in access to care among patients with premature acute coronary syndrome. *Canadian Medical Association Journal*. 2014;186(7):497-504. doi: 10.1503/cmaj.131450.

32. Santomauro DF, Mantilla Herrera AM, Shadid J, et al. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *The Lancet*. 2021;398(10312):1700-1712. doi: 10.1016/S0140-6736(21)02143-7.
33. Douaud G, Lee S, Alfaro-Almagro F, et al. SARS-CoV-2 is associated with changes in brain structure in UK biobank. *Nature (London)*. 2022;604:697-707. doi: 10.1038/s41586-022-04569-5.
34. Al-Aly Z, Bowe B, Xie Y. Long COVID after breakthrough SARS-CoV-2 infection. *Nature Medicine*. 2022. doi: 10.1038/s41591-022-01840-0.

Appendices

Appendix A. Somatic diseases included in as chronic disease

Overarching disease type	Examples hereof given in Lifelines questionnaire[a]
Cardiovascular disease	High blood pressure
	Heart attack
	Narrowing of the arteries in the legs
	Stroke/TIA
	Other heart and/or coronary diseases
Lung disease	Asthma
	COPD
	Chronic bronchitis
Liver disease	Cirrhosis
Kidney disease	Reduced kidney function
Diabetes	Diabetes mellitus Type 1
	Diabetes mellitus Type 2
Chronic muscle disease	MS
Auto-immune illness	Celiac disease
	Inflammatory bowel disorder
	Rheumatoid Arthritis
	Lupus
Cancer	Any form of cancer
Neurological disease	Dementia
	Parkinson's disease
	Alzheimer's disease
Problems with the spleen	Sickle cell anemia
	Removal of spleen
Other chronic health conditions	Open answer option

[a]These are examples given in the Lifelines questionnaires, participants with chronic diseases beyond these examples were able to mention these as well.

Appendix B. Baseline comparison of COVID-19-positive population and the COVID-19-negative population

		COVID-19-positive participants (n=4,231)	Controls (n=8,462)
Completed questionnaires, N		62,224	140,810
Completed questionnaires per participant, median (IQR)		17 (8-23)	20 (12-24)
Female sex, N (%)		2,779 (65.7)	5,558 (65.7)
Age, mean (SD)		52.4 (11.7)	54.0 (12.4)
Educational level, N (%)	Low	467 (11.0)	915 (10.8)
	Medium	2,178 (51.5)	4,144 (49.0)
	High	1,428 (33.8)	3,215 (38.0)
Chronic disease, N (%)[a]	Absent	3,434 (81.2)	7,332 (86.6)
	Present	397 (9.4)	617 (7.3)
Smoking, N (%)	No	3,899 (92.2)	7,721 (91.2)
	Yes	206 (4.9)	603 (7.1)
Method of diagnosis, N (%)	Physician's diagnosis	899 (21.2)	n.a.
	Positive PCR test	3332 (78.8)	n.a.
Hospitalization, N (%)		142 (2.7)	n.a.

[a]See Appendix A for the full list of included chronic diseases.

Appendix C. Completed questionnaires by COVID-19-positive participants stratified by sex and timeframe

Days surrounding the COVID-19 diagnosis	Completed surveys by male COVID-19-positive participants, N (%)	Completed surveys by female COVID-19-positive participants, N (%)
< -401	367 (1.7)	660 (1.6)
-400 - -350	1,163 (5.3)	1,946 (4.8)
-349 - -300	1,754 (8.0)	2,996 (7.4)
-299 - -250	2,065 (9.4)	3,564 (8.9)
-249 - -200	2,188 (10.0)	3,784 (9.4)
-199 - -150	2,154 (9.8)	3,796 (9.4)
-149 - -100	1,733 (7.9)	3,011 (7.5)
-99 - -50	1,470 (6.7)	2,578 (6.4)
-49 – 0	2,815 (12.8)	5,277 (13.1)
1 – 10	226 (1.0)	457 (1.1)
11 – 20	368 (1.7)	781 (1.9)
21 – 30	535 (2.4)	1,077 (2.7)
31 – 40	336 (1.5)	620 (1.5)
41 – 50	376 (1.7)	711 (1.8)
51 – 60	311 (1.4)	655 (1.6)
61 – 70	242 (1.1)	484 (1.2)
71 – 80	292 (1.3)	516 (1.3)
81 – 90	250 (1.1)	517 (1.3)
91 – 100	241 (1.1)	447 (1.1)
101 – 110	246 (1.1)	440 (1.1)
111 – 120	144 (0.7)	275 (0.7)
121 – 150	557 (2.5)	1,123 (2.8)
151 – 200	658 (3.0)	1,356 (3.4)
201 – 250	636 (2.9)	1,302 (3.2)
251 – 300	265 (1.2)	618 (1.5)
301 – 350	175 (0.8)	416 (1.0)
351 – 400	217 (1.0)	441 (1.1)
401 >	191 (0.9)	401 (1.0)
Total	21,975 (100%)	40,249 (100%)

Appendix D. Mean symptom sumscore over time in the uninfected control population

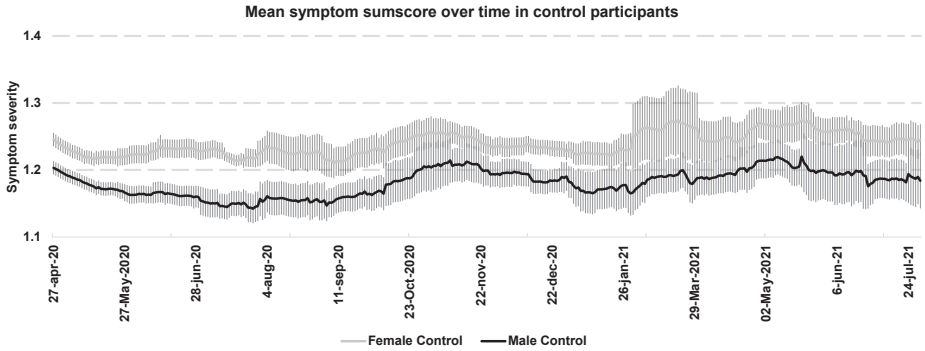


Figure D: The mean somatic symptom sumscore, depicted over time in the uninfected and matched control population.

Appendix E: Sensitivity analyses of symptom prevalence, stratified by COVID-19 diagnosis method.

Symptom	Presence of symptom		
	Controls (n=4353)	COVID-19 (diagnosis and positive test; n=1942)	COVID-19 (positive test; n=1592)
Acute phase symptoms			
Ageusia/anosmia	37 (0.8)	158 (8.1)*	150 (9.4)
Difficulties when breathing	38 (0.9)	68 (3.5)*	51 (3.2)
Chest pain	44 (1.0)	63 (3.2)	49 (3.0)
Pain when breathing	13 (0.3)	20 (1.0)	14 (0.9)
Lump in throat	59 (1.4)	61 (3.1)	49 (3.1)
Heavy arms/legs	130 (3.0)	126 (6.5)	108 (6.8)
General tiredness	159 (3.7)	136 (7.0)	113 (7.1)
Painful muscles	378 (8.7)	262 (13.5)	214 (13.4)
Tingling extremities	145 (3.3)	98 (5.0)	85 (5.3)
Fever	19 (0.4)	16 (0.8)	12 (0.8)
Wet cough	83 (1.9)	58 (3.0)	47 (3.0)
Dry cough	81 (1.9)	50 (2.6)	39 (2.4)
Headache	239 (5.5)	166 (8.5)	133 (8.4)
Itchy eyes	143 (3.3)	96 (4.9)	79 (5.0)
Alternately feeling hot/cold	155 (3.6)	112 (5.8)	90 (5.7)
Sore throat	84 (1.9)	48 (2.5)	38 (2.4)
Runny nose	217 (5.0)	110 (5.7)	87 (5.5)
Nausea	128 (2.9)	72 (3.7)	54 (3.3)
Sneezing	210 (4.8)	101 (5.2)	81 (5.1)
Back pain	413 (9.5)	210 (10.8)	172 (10.5)
Stomach pain	108 (2.5)	53 (2.7)	37 (2.3)
Dizziness	93 (2.1)	46 (2.4)	39 (2.4)
Diarrhea	80 (1.8)	38 (2.0)	27 (1.7)
Total	1,275 (29.3%)	790 (40.7)	642 (40.3)

Data are n (%). Symptoms are ordered according to their relative increase in frequency in COVID-19-positive participants compared with controls. A substantial increase in severity was defined as an increase in symptom severity of at least 1 point on the 5-point scale. †n=3988, sneezing was assessed in 23 surveys instead of 24; ‡n=1704; §n=1536 (100%).

Substantial increase of symptom severity to at least moderate severity		
Controls (n=4130)	**COVID-19 (diagnosis and positive test; n=1782)**	**COVID-19 (positive test; n=1551)**
17 (0.4)	135 (7.6)	133 (8.6)
21 (0.5)	43 (2.4)	38 (2.5)
24 (0.6)	43 (2.4)	37 (2.4)
<10 (<0.1)	16 (0.9)	12 (0.8)
24 (0.6)	42 (2.4)	41 (2.6)
65 (1.6)	75 (4.2)	71 (4.6)
87 (2.1)	88 (4.9)	82 (5.3)
134 (3.2)	130 (7.3)	119 (7.7)
65 (1.6)	52 (2.9)	46 (3.0)
18 (0.4)	12 (0.7)	10 (0.6)
40 (1.0)	28 (1.6)	26 (1.7)
43 (1.0)	28 (1.6)	25 (1.6)
111 (2.7)	76 (4.3)	70 (4.5)
78 (1.9)	51 (2.9)	49 (3.2)
70 (1.7)	63 (2.5)	58 (3.7)
51 (1.2)	29 (1.6)	26 (1.7)
94 (2.3)	50 (2.8)	41 (2.6)
74 (1.8)	37 (2.1)	29 (1.9)
74 (1.9)[†]	35 (2.1)[‡]	32 (2.1)[§]
182 (4.4)	88 (4.9)	77 (5.0)
58 (1.4)	25 (1.4)	19 (1.2)
56 (1.4)	25 (1.4)	24 (1.5)
52 (1.3)	19 (1.1)	17 (1.1)
749 (18.1)	526 (29.6)	447 (28.8)

PART 3

Primary care help-seeking for common somatic symptoms

CHAPTER 7

Ballering, A.V., Olde Hartman, T.C., Verheij, R., & Rosmalen, J.G.M. (2023). Sex and gender differences in primary care help-seeking for common somatic symptoms: a longitudinal study. *Scandinavian Journal of Primary Health Care*, 41(2), 132–139

Abstract

Objective: Women are reported to consult general practitioners (GPs) more frequently than men. However, previous studies on sex differences in help-seeking behaviour for somatic symptoms do not distinguish between sex and gender, do not account for sex differences in presented symptoms, and are frequently conducted in clinical settings, automatically excluding non-help seekers. Therefore, we aim to assess the independent associations of sex and gender with primary care help-seeking for somatic symptoms in the general population.

Design and setting: Records from the longitudinal population-based Lifelines Cohort Study were linked to routine electronic health records from GPs.

Subjects: Participants reporting new-onset common somatic symptoms.

Main outcome measures: Associations between sex and gender, operationalized via a novel gender-index, with primary care help-seeking for somatic symptoms and differences in the strength of the association between gender and help-seeking for somatic symptoms between women and men.

Results: Of 20,187 individuals with linked data, 8,325 participants (67.5% female; mean age=44.5 years [SD=12.9]) reported at least one new-onset somatic symptom. Hereof, 255 (3.1%) consulted the GP within 6 weeks of symptom onset. Female sex was positively associated with consulting the GP (OR=1.78; 95%CI=1.13-2.80), whereas feminine gender was not (OR=0.67; 95%CI=0.39-1.16). The latter association did not differ in strength between men and women. More paid working days negatively associated with help-seeking (OR=0.95; 95%CI=0.91-0.98).

Conclusion: The results suggest that female sex rather than feminine gender is associated with primary care help-seeking behaviour for somatic symptoms. Nevertheless, clinicians should be aware that gender-related variables, such as mean paid working days, may associate with help-seeking behaviour.

Introduction

Previous studies have suggested that women more frequently consult their general practitioner (GP) for common somatic symptoms than men.[1-5] This may partly be explained by sex and gender differences in the occurrence of such symptoms: female participants reported more frequent, severe and persistent somatic symptoms, whereas femininity (i.e. feminine gender) was protective of the persistence of symptoms.[6-8] However, also help-seeking behavior for these symptoms might contribute to the increased consultation rates in women. To understand how help-seeking for somatic symptoms is affected by sex and gender, it is important to clearly distinguish between these latter two concepts.[9] Sex refers to biological characteristics, including but not limited to chromosomes, hormones, and anatomy, of female and male bodies. In contrast, gender is a socioculturally-constructed, multidimensional concept that entails the embodiment of different roles, behaviors, identities, and relationships of women and men prescribed by social norms in a given time and society.[10]

The idea of increased female primary care help-seeking for common somatic symptoms, however, is under debate. A recent systematic review that directly compared the consultation patterns in women and men experiencing headaches and lower back pain concluded that the evidence for a female preponderance in help-seeking was surprisingly weak and inconsistent.[11] Previous studies on sex and gender differences in primary care help-seeking for somatic symptoms are characterized by several methodological limitations.

First, potentially due to the absence of adequate gender measures in epidemiological studies, these could not quantify independent sex and gender differences in the frequency of help-seeking behavior for somatic symptoms.[6,9,12-14] Second, most studies focusing on help-seeking behavior in relation to somatic symptoms are conducted in clinical populations or patient registries, which is problematic as it automatically excludes people who do not seek help for their somatic symptoms.[11] Third, previous studies are largely based on self-reported measures of help-seeking, making these prone to recall bias. Fourth, most study designs do not correct for the type of somatic symptom and sex differences in the presentation of these.[15]

We present the first large epidemiological cohort study on the association between sex, gender and help-seeking for common somatic symptoms that overcomes these problems. We linked the Lifelines general population cohort to the Nivel Primary Care Database (NPCD),[16] allowing us to assess independent sex and gender differences in symptom-specific help-seeking behavior in the general population. We examined

whether sex and gender have a unique association with help-seeking behavior for common somatic symptoms and whether the association between gender and help-seeking behavior for common somatic symptoms varies between sexes.

Methods

Lifelines and the Nivel Primary Care Database

Lifelines is a multi-disciplinary prospective population-based cohort study examining in a unique three-generation design the health and health-related behaviors of 167,729 persons living in the North of The Netherlands. It employs a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioral, physical and psychological factors which contribute to health and disease of the general population, with a special focus on multi-morbidity and complex genetics. The Lifelines Cohort Study is approved by the Medical Ethical Committee of University Medical Center Groningen (2007/152) and participants provided written consent. Extensive information on the cohort, design considerations and recruitment procedures is provided elsewhere.[17,18] In this study, we included Lifelines data collected from 2008 up until 2018.

The NPCD encompasses electronic health records data from general practices located across the Netherlands, with diagnoses, contacts and prescription medicine coded via the International Classification of Primary Care (ICPC-2) system. Virtually all people residing in the Netherlands are listed as a patient in a GP practice. The NPCD population is a representative sample of the Dutch population. Data on GP consultations for common somatic symptoms as described in the Symptom CheckList-90 Somatization subscale (SCL-90 SOM; ***Appendix A)*** were retrieved from electronic health records for all participants that were listed in one of the 63 NPCD GP practices located in the North of the Netherlands. This study has been approved according to the governance code of NPCD under number NZR-00319.049.

Variables

The presence of self-reported common somatic symptoms was assessed in Lifelines surveys by the 12-item ordinal SCL-90 SOM, which has been recommended for large-scale studies and has sufficient measurement invariance over time.[19,20] ***Appendix B*** shows the definition of new-onset symptoms. The symptoms (ICPC) of the SCL-90 SOM comprise headache (N01), dizziness (N17), heart pain (K01), (lower) back pain (L02 and L03), nausea (D09), muscle pain (L18), shortness of breath (R02), chills (A02), tingling of fingers, feet and/or toes (N05), swallowing/throat problems (D21 and R21),

general tiredness (A04), and heavy arms and/or legs (L09 and L14). The SCL-90 SOM subscale assesses the extent of distress or bother participants experienced during the past seven days due to these symptoms. Presence of the new-onset symptoms was assessed with Lifelines' survey data, whereas GP consults for these symptoms were listed in the NPCD.

Participants' sex (female or male) and age in years were derived from the municipal population registry. Participants' self-reported highest attained educational level was categorized into 'high', 'medium' and 'low'.[21] Participants' burden of somatic symptoms was measured by the SCL-90 SOM sumscore at the moment of reporting a new-onset symptom.[22] Participants' feminine and masculine gender roles were operationalized via a recently developed, data-driven gender index, which accounts for the place-, time- and society-bound nature of gender roles.[6] In a subsample of baseline Lifelines adult participants that had no suspected intersex variation or gender-diverse gender identity, we performed LASSO logistic regression analyses that used 153 psychosocial characteristics, including dietary preferences, hobbies, time spend on household tasks or odd jobs, type of profession and personality traits, to predict participants' municipally-registered sex. The majority of the included psychosocial variables referred to gender roles. Using the obtained estimates of the regression coefficients of the LASSO logistic regression model (AUC=92%), an individual score regarding feminine and masculine gender roles (i.e. the gender index) was calculated for each adult participant. In other words, participants' individual adherence to prototypical feminine and masculine psychosocial characteristics was calculated. The gender index ranges from 0% (fully masculine) to 100% (fully feminine). An index of 50% indicates androgyny, with equal levels of feminine and masculine characteristics present. Extensive information on the development and applicability of the gender index is provided elsewhere.[6,7]

To define help-seeking behavior of participants we linked data derived from Lifelines with the NPCD (for detailed procedure: **Appendix C**). A Dutch trusted third party (Statistics Netherlands) used Record Identification Number pseudonymization to temporarily link health records on an individual level to facilitate analyses.[23] The presence of help-seeking was defined as the presence of a GP contact (either face-to-face, by phone or digital) for the reported symptom provided 6 weeks before, or 6 weeks after reporting a new-onset common somatic symptom during a Lifelines assessment. Presence of help-seeking within 3 months of symptom reported was defined using a similar strategy. Since we only included contacts with ICPC≤30 codes, we restricted our analyses to help-seeking for common somatic symptoms that the GP did not diagnose with an underlying disease (i.e. disease diagnosis; ICPC≥70), but with a symptom diagnosis (i.e. symptoms that remain symptoms over time).[24]

Statistical analyses

To assess whether sex and gender were independently associated with primary care help-seeking behavior we applied generalized linear mixed-effect models with maximum likelihood estimation. This approach allows for accounting for the dependency of residual errors due to the hierarchical structure of the data; these models also handle missing data efficiently. Our data were clustered on three levels: (1) observations of help-seeking (2) were nested within patients, (3) who in turn were nested in GP practices. Our initial model included only the intercept as independent variable and allowed intercepts to vary across individuals and GP practices (i.e. random intercepts for the second and third level). One-by-one we included independent variables as fixed effects. Thereafter we also allowed the effects of sex and gender to vary across patients and GP practices (i.e. random intercepts and slopes for the second and third level). Model fit was assessed using the Akaike Information Criterion (AIC), with a lower AIC indicating better model fit. One-way ANOVAs were applied to assess the significance of differences in model fit. The random effects' covariance matrix was unstructured and no mean centering of continuous independent variables was performed, as these had meaningful zero points.

Included independent variables were participants' sex, age, educational level, score on the gender index, reported new-onset symptom, burden of somatic symptom experience at the moment of symptom reporting, and self-reported lifetime presence of chronic somatic and psychiatric diseases at the moment of symptom reporting. We also included a sex-by-gender interaction term to assess whether the association between gender and help-seeking differed for female and male participants. We repeated these analyses with help-seeking within 3 months as dependent variable.

In post-hoc analyses we assessed whether gender-related factors are associated with help-seeking. We repeated the abovementioned analyses, but replaced the gender index with gender-related factors, namely being a healthcare professional (yes/no), mean days per week one performs paid labor (1-7 days) and whether one considers oneself a homemaker (yes/no) as these were indicated by previous qualitative research to be of importance for help-seeking.[9,13,14]

We assessed the presence of multicollinearity among independent variables, but found no indication of problems with multicollinearity as the variance inflation factor was ≤ 5 in all analyses.[25] We adhered to a two-sided α-level of 0.003, corrected for multiple comparisons (0.05/20; 19 independent variables and one interaction term within the family of tests). Data on independent variables were imputed as described earlier.[6] All analyses were conducted using the *lme4* package in R version 3.6.2 (R Foundation for Statistical Computing, Vienna, Austria).

Results

Patient population and GP consults

The Lifelines baseline population consisted of 152,728 adult participants, whereas the NPCD population comprised 277,881 patients. In total, we linked 20,187 individuals of whom 2709 (32.5%) male and 5616 (67.5%) female participants reported new-onset symptoms (***Table 1***).

Table 1. Characteristics of the study population at baseline (N=8,325)

		Male participants (N=2,709; 32.6%)	**Female participants (N=5,616; 67.5%)**
Age in years, mean (SD)		44.8 (12.7)	42.9 (13.0)
Gender index, median (IQR)		0.08 (0.02-0.30)	0.95 (0.82-0.99)
Educational level, N (%)	Low	905 (33.4%)	1,724 (30.7%)
	Medium	968 (35.7%)	2,229 (39.7%)
	High	836 (30.9%)	1,663 (29.6%)
Symptom burden, mean SCL-90 SOM score (SD)		1.38 (0.37)	1.49 (0.41)
Chronic disease, N (%)	Present	1,241 (45.8%)	3,035 (54.0%)
	Absent	1,468 (54.2%)	2,581 (46.0%)
Self-rated health, N (%)	Good to excellent	2,326 (87.1%)	4,728 (85.4%)
	Poor to mediocre	345 (12.9%)	807 (14.6%)
New-onset symptom		5,439	12,416
Help-seeking 6 weeks		70	185
Help-seeking 3 months		111	278

Ultimately, 255 individual participants (3.1%) sought help 360 times for 255 different symptoms within 6 weeks of reporting a new-onset symptom. Within 3 months of reporting a new-onset symptom, 387 individual participants (4.6%) sought help 596 times for 389 different new-onset symptoms (***Table 2***). The most frequently reported new-onset symptoms are muscle pains and (lower) back pain in 1037 (38.3%) and 903 (33.3%) of the male and in 1979 (35.2%) and 1626 (29.0%) of the female participants with new-onset symptoms, respectively (***Appendix 4***).

Table 2. Reported new-onset symptoms in Lifelines and concomitant GP contacts in 8,325 participants.

New-onset common somatic symptom (ICPC)	Symptoms reported, N (%)	GP contacts within 6 weeks, N (%)	GP contacts within 3 months, N (%)
Headache (N01)	1,868 (10.5%)	19 (7.5%)	26 (6.7%)
Dizziness (N17)	761 (4.3%)	17 (6.7%)	24 (6.2%)
Heartpain (K01)	483 (2.7%)	<10 (<3.9%)	<10 (2.6%)
(Lower) backpain (L02/L03)	2,529 (14.2%)	88 (34.5%)	138 (35.5%)
Nausea (D09)	1,406 (7.9%)	<10 (<3.9%)	<10 (<2.6%)
Muscle pain (L18)	3,016 (16.9%)	20 (7.8%)	39 (10.0%)
Shortness of breath (R02)	633 (3.5%)	13 (5.1%)	17 (4.4%)
Hot-and-cold spells (A02)	1,627 (9.1%)	<10 (<3.9%)	<10 (<2.6%)
Tingling extremities (N05)	1,566 (8.7%)	10 (3.9%)	12 (3.1%)
Throat/swallowing problems (D21/R21)	798 (4.5%)	17 (6.7%)	25 (6.4%)
General tiredness (A04)	1,698 (9.5%)	43 (16.9%)	66 (17.0%)
Arm/leg symptoms (L09/L14)	1,470 (8.2%)	19 (7.5%)	28 (7.2%)
Total	17,855 (100.0%)	255 (100.0%)	389 (100.0%)

Sex and gender in association with help-seeking behavior for somatic symptoms

We defined generalized linear mixed-effect models with increasing complexity to assess the associations between sex and gender, and help-seeking. For help-seeking within 6 weeks of new-onset symptom reporting, the model that allowed the effect of only sex (AIC = 2495.1), or of both sex and gender to vary across GP practices (AIC =2501.1; i.e. models including random slopes) did not fit the data significantly better than the model that included only random intercepts for patients and GP practices (AIC = 2491.6; $\chi^2_{(DF=2)}$=0.58, p=0.75 and $\chi^2_{(DF=5)}$=0.58, p=0.99). Similarly, for the 3 month timeframe, the model that included random intercepts for patients and GP practices (AIC = 3454.4) fitted the data better than the models that allowed sex to vary across GP practices (AIC = 3458.1), or sex and gender to vary across GP practices (AIC=3464.0). However, the differences in model fit were not statistically significant ($\chi^2_{(DF=2)}$=0.37, p=0.83 and $\chi^2_{(DF=5)}$=0.41, p=0.99, respectively). Therefore, we present the models including random intercepts, as these had the lowest AIC (***Table 3***).

The associations between the independent variables and seeking help within 6 weeks of symptom reporting are similar in effect size to those of help-seeking within 3 months of symptom reporting. Female patients seek help more often within 6 weeks of new-onset symptom reporting than male patients (OR=1.78, 95%CI=1.13-

2.80), but femininity was not associated with help-seeking (OR=0.67, 95%CI=0.39-1.16). The sex-by-gender interaction terms were not statistically significant (OR = 0.54, 95%CI=0.18-1.65), indicating that the association between gender and help-seeking within 6 weeks does not differ between female and male patients.

Table 3. Generalized linear mixed-effects models: estimated associations between independent variables and help-seeking within different time frames.

		Help-seeking (N=17,855)	
		Within 6 weeks	**Within 3 months**
Fixed effects		**OR (95%CI)**	
Female sex		1.78 (1.13-2.80)*	1.53 (1.05-2.21)*
Femininity		0.67 (0.39-1.16)	0.76 (0.49-1.20)
Age		1.03 (1.02-1.04)*	1.02 (1.01-1.03)*
Presence of chronic disease		1.00 (0.76-1.31)	1.04 (0.84-1.30)
Educational level	Low	Ref.	Ref.
	Medium	1.22 (0.89-1.67)	0.95 (0.74-1.23)
	High	1.18 (0.84-1.66)	0.94 (0.71-1.24)
Symptom burden		0.98 (0.96-1.01)	0.99 (0.97-1.01)
Presence of symptom	Headache	0.78 (0.41-1.48)	0.74 (0.43-1.28)
	Dizziness	1.77 (0.92-3.41)	1.67 (0.97-2.90)
	Heartpain	0.47 (0.14-1.58)	0.54 (0.21-1.38)
	(Lower) backpain	2.67 (1.62-4.40)*	3.00 (1.99-4.53)*
	Nausea	0.29 (0.11-0.76)*	0.31 (0.14-0.68)*
	Muscle pain	0.48 (0.26-0.90)*	0.65 (0.40-1.06)
	Shortness of breath	1.57 (0.77-3.19)	1.40 (0.76-2.56)
	Hot-and-cold spells	0.05 (0.01-0.32)*	0.03 (0.01-0.22)*
	Tingling extremities	0.47 (0.28-0.99)*	0.38 (0.19-0.75)*
	Throat problems	1.67 (0.87-3.21)	1.68 (0.98-2.90)
	General tiredness	2.10 (1.22-3.59)*	2.24 (1.43-3.49)*
	Heavy arms/legs	1.27 (0.68-2.41)	1.33 (0.78-2.28)

*p<0.003

Post-hoc analyses

Femininity operationalized by the gender index showed no association with help-seeking for new-onset common somatic symptoms. We assessed whether specific gender-related factors that were identified by previous qualitative research were associated with help-seeking. We found that working as a healthcare professional or considering oneself as a homemaker showed no association with help-seeking within

6 weeks (OR=0.93; 95%CI=0.65-1.32 and OR=0.68; 95%CI=0.42-1.14, respectively). Weekly mean days of paid work did associate with help-seeking within 6 weeks (OR=0.95; 95%CI=0.91-0.98). No associations were found with help-seeking within 3 months.

Discussion

To our knowledge, this is the first large-scale study that assesses separate associations of sex and gender with primary care help-seeking behavior for common somatic symptoms. Female sex was associated with help-seeking within 6 weeks and 3 months of symptom reporting, whereas femininity was not. We found increased mean working days per week to be negatively associated with help-seeking behavior.

Strengths and limitations

This study had several strengths. First, we directly compared help-seeking in women and men adjusted for the reported symptoms. Previous studies did not adjust for the reported symptoms and sex differences in presentation hereof. Second, previous studies assessing differences in primary care help-seeking behavior did not distinguish between sex and gender.[11,26] Here, we used a novel gender measure to disentangle associations of sex and gender with help-seeking behavior. Third, this study included consultation patterns based on primary care registries as opposed to self-reported measures, resulting in minimal risk of recall bias. Lastly, this study assessed help-seeking in the general population instead of clinical populations, which allowed for participants who have not sought help to be included in the analyses.

However, this study also had limitations. First, the NPCD is physician-centered and diagnosis-based, and the reason for encounter as reported by patients was not recorded. The use of recorded final diagnoses in this study implies that the frequency of help-seeking behavior was underestimated: GP consults in which a disease was recorded (i.e. ICPC≥70) were excluded from analyses while those for which a symptom diagnosis was recorded were included (i.e. ICPC≤30). This may explain the negative association between nausea, muscle pain, hot-and-cold spells and tingling extremities with GP contacts, as these are hardly diagnosed with symptom diagnoses. A recent study showed that men had a 6%-increase in odds of being provided with a disease diagnosis for their somatic symptoms compared to women.[27] This may imply that male help-seeking is underestimated in this study, and that the reported sex difference is overestimated.

Second, the exact moment of symptom onset is unknown and symptoms could have been present for a substantial amount of time before individuals decide to consult their GP. Possibly, patients already sought help for symptoms in an earlier stage which would erroneously classify these participants as non-help seekers. This may also lead to an underestimation of help-seeking behavior, potentially explaining why a mere 3.1% of the patients with new-onset symptoms sought help, which is less than reported by previous studies.[28] Moreover, many previous studies assessing primary care help-seeking behavior include symptoms beyond the 12 common somatic symptoms analyzed in this study, including sex-related symptoms and symptoms that require acute help-seeking, such as traumatic injuries, and do not distinguish between first and follow-up visits.[2-4] Although no studies about sex differences in the amount of time between the moment of first symptoms until the moment of first contact in primary care for common somatic symptoms are known to the authors, studies on symptoms associated with cancer and stroke among 10,297 and 162,856 adult patients, respectively, report no or inconclusive sex differences in time to help-seeking.[29,30] This suggests that no sex bias was introduced due to delays in help-seeking in this study.

Comparison with previous studies

We found that female patients sought help more frequently from their GP for common somatic symptoms than male patients. This is in line with previous studies,[1,2,8,31] including those that adjusted for sex-specific symptoms.[5,15] Other studies stated that the evidence for a sex difference in help-seeking for common somatic symptoms is weak.[11,26] However, these studies did not solely focus on primary care, included other symptoms than the current study, defined symptom experience very broadly, or relied on self-reported data about help-seeking behavior. These differences in methodology may result in differing outcomes between the studies.

Often, sex differences in help-seeking are, at least partly, attributed to gender differences between women and men.[15] However, the results of this study suggest that factors associated with sex differences in help-seeking should be sought in either the biological realm or in factors that go beyond the composite gender index. Multiple reasons grounded in biology have been explored for the female preponderance in help-seeking behavior. First, women experience, describe and report their symptoms in a different manner, and more readily attribute these to somatic causes than men.[8,32] Second, depressive and anxiety-related symptoms are more prevalent in women than in men, and these depressive and anxiety-related symptoms strongly associate with help-seeking.[5,8] Analyses in this study were adjusted for diagnosed psychiatric disorders, but not for depressive or anxiety-related symptoms. Last, female patients

may have a lower threshold to seek help from their GP, as they are more familiar with primary care. This familiarity may arise, for example, because of more frequent visits to the GP related to gynaecological interventions (i.e. the recurring smear test or pregnancy-related visits).

Gender-related factors that are not incorporated or are diluted in the gender index may also affect help-seeking. The gender index is based on psychosocial variables that predominantly reflect gender roles. Although studies show that more practical gender-related factors, such as household responsibilities,[33,34] influence help-seeking behavior, others argue that mainly gender stereotypes influence people's help-seeking behavior.[35,36] Gender stereotypes are not incorporated in the gender index, yet these do pose a social framework on people resulting in gendered behaviors and ideas. For example, traditional western gender stereotypes prescribe that it is more culturally and socially accepted for women to openly express their symptoms compared to men.[6] In contrast, these gender stereotypes state that men should be stoical about bodily experiences and should conceal their symptoms, even from care providers.[36,37] Although the gender index does not associate with help-seeking behavior, gender as an influencing factor on help-seeking should not be discarded completely.

Clinical implications

This study found that female sex and the amount of days performing paid labor both associate with help-seeking behavior. For clinicians the patient's characteristics, such as sex, frequency of help-seeking and occupational factors, are pivotal in the clinical decision-making processes. Therefore, it is important for clinicians to be aware of female sex and gender-related factors being associated with primary care help-seeking[38]. Gender stereotyping, as mentioned above for example, may impose a social framework on people affecting help-seeking behavior. Awareness hereof may also counter patients' delayed help-seeking behavior, which may result in delayed detection and concomitant treatment of symptoms.

Further research should consider whether the sex difference in help-seeking behavior also results in a sex difference in conducted diagnostic tests and diagnosed diseases.[27,39] Men are often typed as more reluctant seekers of healthcare and health information, whereas women are portrayed as frequent help-seekers. Such ideas may inadvertently prompt the GP to consider men's symptoms as more serious than women's symptoms, resulting in less watchful waiting in men.

References

1. Rodriguez González AM, Ramírez Aranda JM, de Los Santos Reséndiz H, et al. Health care utilization of mexican patients with medically unexplained physical symptoms. *Colombia Médica*. 2016;47(3):155-159.
2. de Waal M, Arnold IA, Eekhof J, Assendelft W, van Hemert AM. Follow-up study on health care use of patients with somatoform, anxiety and depressive disorders in primary care. *BMC Family Practice*. 2008;9:5. doi: 10.1186/1471-2296-9-5.
3. Redondo-Sendino Á, Guallar-Castillón P, Banegas JR, Rodríguez-Artalejo F. Gender differences in the utilization of health-care services among the older adult population of spain. *BMC Public Health*. 2006;6(1):155. doi: 10.1186/1471-2458-6-155.
4. Bertakis KD, Azari R, Helms LJ, Callahan EJ, Robbins JA. Gender differences in the utilization of health care services. *Journal of Family Practice*. 2000;49(2):147.
5. Wang Y, Hunt K, Nazareth I, Freemantle N, Petersen I. Do men consult less than women? an analysis of routinely collected UK general practice data. *BMJ Open*. 2013;3(8):e003320-003320.
6. Ballering AV, Bonvanie IJ, Olde Hartman TC, Monden R, Rosmalen JGM. Gender and sex independently associate with common somatic symptoms and lifetime prevalence of chronic disease. *Social Science & Medicine*. 2020;253:112968. doi: 10.1016/j.socscimed.2020.112968.
7. Ballering AV, Wardenaar KJ, olde Hartman TC, Rosmalen JGM. Female sex and femininity independently associate with common somatic symptom trajectories. *Psychological Medicine*. 2022;52(11):2144-2154. doi: 10.1017/S0033291720004043.
8. Barsky AJ, Peekna HM, Borus JF. Somatic symptom reporting in women and men. *Journal of General Internal Medicine*. 2001;16(4):266-275. doi: 10.1046/j.1525-1497.2001.00229.x.
9. Doyal L. Sex and gender: The challenges for epidemiologists. *International Journal of Health Services*. 2003;33(3):569-579.
10. Johnson JL, Greaves L, Repta R. Better science with sex and gender: Facilitating the use of a sex and gender-based analysis in health research. *International Journal for Equity in Health*. 2009;8(1):14. doi: 10.1186/1475-9276-8-14.
11. Hunt K, Adamson J, Hewitt C, Nazareth I. Do women consult more than men? A review of gender and consultation for back pain and headache. *Journal of Health Services Research & Policy*. 2011;16(2):108-117. doi: 10.1258/jhsrp.2010.009131.
12. Nowatzki N, Grant KR. Sex is not enough: The need for gender-based analysis in health research. *Health Care for Women International*. 2011;32(4):263-277. doi: 10.1080/07399332.2010.519838.
13. Currie D, Wiesenberg S. Promoting women's health-seeking behavior: Research and the empowerment of women. *Health Care for Women International*. 2003;24(10):880-899. doi: 10.1080/07399330390244257.
14. Tenenbaum A, Nordeman L, Sunnerhagen KS, Gunnarsson R. Gender differences in care-seeking behavior and healthcare consumption immediately after whiplash trauma. *PloS one*. 2017;12(4):e0176328. doi: 10.1371/journal.pone.0176328.
15. Hunt K, Adamson J, Galdas P. Gender and help-seeking: Towards gender-comparative studies. In: *The palgrave handbook of gender and healthcare*. London: Palgrave Macmillan UK; 2010:207-221. 10.1057/9780230290334_13.
16. Hasselaar J, Overbeek L. Nivel zorgregistraties eerste lijn - databron ziekten en zorg in de eerste lijn. https://www.nivel.nl/nl/nivel-zorgregistraties-eerste-lijn. Updated 2023. Accessed 1 June, 2023.

17. Scholtens S, Smidt N, Swertz MA, et al. Cohort profile: LifeLines, a three-generation cohort study and biobank. *International Journal of Epidemiology*. 2015;44(4):1172-80. doi: 10.1093/ije/dyu229.
18. Klijs B, Scholtens S, Mandemakers JJ, Snieder H, Stolk RP, Smidt N. Representativeness of the LifeLines cohort study. *PLoS ONE*. 2015;10(9):e0137203. doi: 10.1371/journal.pone.0137203.
19. Rytilä-Manninen M, Fröjd S, Haravuori H, et al. Psychometric properties of the symptom checklist-90 in adolescent psychiatric inpatients and age- and gender-matched community youth. *Child and Adolescent Psychiatry and Mental Health*. 2016;10(23):23. doi: 10.1186/s13034-016-0111-x.
20. Zijlema WL, Stolk RP, Löwe B, Rief W, White PD, Rosmalen JGM. How to assess common somatic symptoms in large-scale studies: A systematic review of questionnaires. *Journal of Psychosomatic Research*. 2013;74(6):459-68. doi: 10.1016/j.jpsychores.2013.03.093.
21. van Zon SKR, Snieder H, Bültmann U, Reijneveld SA. The interaction of socioeconomic position and type 2 diabetes mellitus family history: A cross-sectional analysis of the lifelines cohort and biobank study. *BMJ Open*. 2017;7(4):e015275. doi: 10.1136/bmjopen-2016-015275.
22. Aaronson NK, Muller M, Cohen PDA, et al. Translation, validation, and norming of the dutch language version of the SF-36 health survey in community and chronic disease populations. *J Clin Epidemiol*. 1998;51(11):1055-1068. doi: 10.1016/S0895-4356(98)00097-3.
23. Kuchinke W, Ohmann C, Verheij RA, et al. A standardised graphic method for describing data privacy frameworks in primary care research using a flexible zone model. *International journal of medical informatics*. 2014;83(12):941-957. doi: 10.1016/j.ijmedinf.2014.08.009.
24. Soler JK, Okkes I. Reasons for encounter and symptom diagnoses: A superior description of patients' problems in contrast to medically unexplained symptoms (MUS). *Family practice*. 2012;29(3):272-282. doi: 10.1093/fampra/cmr101.
25. Miles J, Shevlin M. Collinearity. In: *Applying regression and correlation: A guide for students and researchers*. Sage; 2001:126-135.
26. Elnegaard S, Andersen RS, Pedersen AF, et al. Self-reported symptoms and healthcare seeking in the general population -exploring "The symptom iceberg". *BMC Public Health*. 2015;15(1):685. doi: 10.1186/s12889-015-2034-5.
27. Ballering AV, Muijres D, Uijen AA, Rosmalen JGM, olde Hartman TC. Sex differences in the trajectories to diagnosis of patients presenting with common somatic symptoms in primary care: An observational cohort study. *Journal of Psychosomatic Research*. 2021;149:110589. doi: 10.1016/j.jpsychores.2021.110589.
28. Green LA, Fryer GE, Yawn BP, Lanier D, Dovey SM. The ecology of medical care revisited. *N Engl J Med*. 2001;344(26):2021-2025. doi: 10.1056/NEJM200106283442611.
29. Potisopha W, Vuckovic KM, DeVon HA, Park CG, Hershberger PE. Sex differences in prehospital delay in patients with acute stroke: A systematic review. *The Journal of cardiovascular nursing*. 2020;35(6):E77-E88. doi: 10.1097/JCN.0000000000000715.
30. Keeble S, Abel GA, Saunders CL, et al. Variation in promptness of presentation among 10,297 patients subsequently diagnosed with one of 18 cancers: Evidence from a national audit of cancer diagnosis in primary care. *International journal of cancer*. 2014;135(5):1220-1228. doi: 10.1002/ijc.28763.
31. Osika Friberg I, Krantz G, Määttä S, Järbrink K. Sex differences in health care consumption in sweden: A register-based cross-sectional study. *Scandinavian Journal of Public Health*. 2016;44(3):264-273. doi: 10.1177/1403494815618843.

32. Gijsbers Van Wijk C, Kolk A. Sex differences in physical symptoms: The contribution of symptom perception theory. *Social Science & Medicine*. 1997;45(2):231-246. doi: 10.1016/s0277-9536(96)00340-1.
33. Norcross W, Ramirez C, Palinkas L. The influence of women on the health care-seeking behavior of men. *The Journal of Family Practice*. 1996;43:475-80.
34. Pelletier R, Humphries KH, Shimony A, et al. Sex-related differences in access to care among patients with premature acute coronary syndrome. *Canadian Medical Association Journal*. 2014;186(7):497-504. doi: 10.1503/cmaj.131450.
35. Juvrud J, Rennels JL. "I don't need help": Gender differences in how gender stereotypes predict help-seeking. *Sex Roles*. 2017;76(1):27-39. doi: 10.1007/s11199-016-0653-7.
36. MacLean A, Hunt K, Smith S, Wyke S. Does gender matter? an analysis of men's and women's accounts of responding to symptoms of lung cancer. *Social Science & Medicine*. 2017;191:134-142. doi: 10.1016/j.socscimed.2017.09.015.
37. Galdas PM, Johnson JL, Percy ME, Ratner PA. Help seeking for cardiac symptoms: Beyond the masculine-feminine binary. *Social Science & Medicine*. 2010;71(1):18-24. https://www.ncbi.nlm.nih.gov/pubmed/20398989 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5142841/. doi: 10.1016/j.socscimed.2010.03.006.
38. Berner AM, Lund J, Saunders CL. Tackling the complexity of gender bias in primary care. *British Journal of General Practice*. 2021;71(708):296-297. doi: 10.3399/bjgp21X716177.
39. Ballering A, Rosmalen J, Olde Hartman T. Differences between women and men are present in the rate of diagnosed diseases after a diagnostic intervention is conducted in primary care. *Journal of the American Board of Family Medicine*. 2022;35(1):73-84. doi: 10.3122/jabfm.2022.01.210289.

Appendices

Appendix A. The SCL-90 SOM subscale

The Symptom CheckList-90 somatization subscale (SCL-90 SOM) is a subscale of the Symptom CheckList-90 (SCL-90). It asks participants to score their degree of bothering or distress due to twelve somatic symptoms in the past week. The total score ranges from 12–60, whereas the mean score of the SCL-90 SOM may range from 1-5.

How much in the past week were you bothered by:		**Not at all**	**A little bit**	**Moderately**	**Quite a bit**	**Extremely**
1	Headache	1	2	3	4	5
2	Dizziness	1	2	3	4	5
3	Chest pain	1	2	3	4	5
4	Lower back pain	1	2	3	4	5
5	Nausea	1	2	3	4	5
6	Painful muscles	1	2	3	4	5
7	Difficulties breathing	1	2	3	4	5
8	Feeling hot and cold alternately	1	2	3	4	5
9	Numbness/tingling in parts of your body	1	2	3	4	5
10	Feeling a lump in your throat	1	2	3	4	5
11	Weakness in body parts	1	2	3	4	5
12	Heavy arms or legs	1	2	3	4	5

Appendix D. Sex-stratified reporting of new-onset symptoms and help-seeking behavior

New-onset common somatic symptom (ICPC)	**Reported symptoms, N (%)**	
	Male (N=2,709)	**Female (N=5,616)**
Headache (N01)	477 (8.8%)	1,391 (11.2%)
Dizziness (N17)	194 (3.6%)	567 (4.6%)
Heartpain (K01)	175 (3.2%)	308 (2.5%)
(Lower) backpain (L02/L03)	903 (16.6%)	1,626 (13.1%)
Nausea (D09)	376 (6.9%)	1,030 (8.3%)
Muscle pain (L18)	1,037 (19.1%)	1,979 (15.9%)
Shortness of breath (R02)	209 (3.8%)	424 (3.4%)
Hot-and-cold spells (A02)	289 (5.3%)	1,338 (10.8%)
Tingling extremities (N05)	527 (9.7%)	1,039 (8.4%)
Throat/swallowing problems (D21/R21)	221 (4.1%)	577 (4.6%)
General tiredness (A04)	571 (10.5%)	1,127 (9.1%)
Arm/leg symptoms (L09/L18)	460 (8.5%)	1,010 (8.1%)
Total	5,439 (100.0%)	12,416 (100.0%)

Appendix B. Definition of new-onset common somatic symptoms

New-onset symptoms were identified as symptoms that were not reported as present at baseline, but were reported as present during a follow-up measurement. Presence of symptoms was based on participants SCL-90 SOM score, as a score ≥3 ('moderately', 'quite a bit' and 'extremely') indicated presence of symptoms, whereas ≤2 ('not at all' and 'a little bit') indicated absence of symptoms.

Baseline	FUP1	FUP2	FUP3	New-onset symptom?
0	0	0	0	-
0	x	0	0	Yes, at FUP1
0	x	x	0	Yes, at FUP1
0	x	x	x	Yes, at FUP1
0	x	0	x	Yes, at FUP1
0	0	x	0	Yes, at FUP2
0	0	x	x	Yes, at FUP2
0	0	0	x	Yes, at FUP3
x	0	0	0	-
x	x	0	0	-
x	x	x	0	-
x	x	x	x	-
x	0	x	0	-
x	0	x	x	-
x	0	0	x	-
x	x	0	x	-

Help-seeking within 6 weeks, N (%)		Help-seeking within 3 months, N (%)	
Male (N=2,709)	Female (N=5,616)	Male (N=2,709)	Female (N=5,616)
< 10 (<14.3%)	16 (8.6%)	<10 (<9.0%)	21 (7.6%)
< 10 (<14.3%)	15 (8.1%)	<10 (<9.0%)	21 (7.6%)
< 10 (<14.3%)	<10 (<5.4%)	<10 (<9.0%)	<10 (<3.6%)
25 (35.7%)	63 (34.1%)	42 (37.8%)	96 (34.5%)
< 10 (<14.3%)	<10 (<5.4%)	<10 (<9.0%)	<10 (3.6%)
< 10 (<14.3%)	15 (8.1%)	11 (9.9%)	28 (10.1%)
< 10 (<14.3%)	<10 (<5.4%)	<10 (<9.0%)	10 (3.6%)
< 10 (<14.3%)	<10 (<5.4%)	<10 (<9.0%)	<10 (<3.6%)
< 10 (<14.3%)	<10 (<5.4%)	<10 (<9.0%)	10 (3.6%)
< 10 (<14.3%)	13 (7.0%)	<10 (<9.0%)	18 (6.5%)
16 (22.9%)	27 (14.6%)	23 (20.7%)	45 (15.5%)
< 10 (<14.3%)	15 (8.1%)	<10 (<9.0%)	21 (7.6%)
70 (100.0%)	185 (100.0%)	111 (100.0%)	278 (100.0%)

Appendix C. Procedure and flowchart of included participants and GP consults

We retrieved all GP consultations related to common somatic symptoms from 2008 to 2018 in practices in the North of the Netherlands from the NPCD that could be linked to adult Lifelines participants. We assessed whether these consultations were associated with new-onset symptoms reported in Lifelines surveys. Consultations with the same ICPC codes as the aforementioned symptoms were regarded as related to the reported symptom. If participants contacted the GP multiple times within the assessed timeframe for the same symptom, we only included the first contacts.

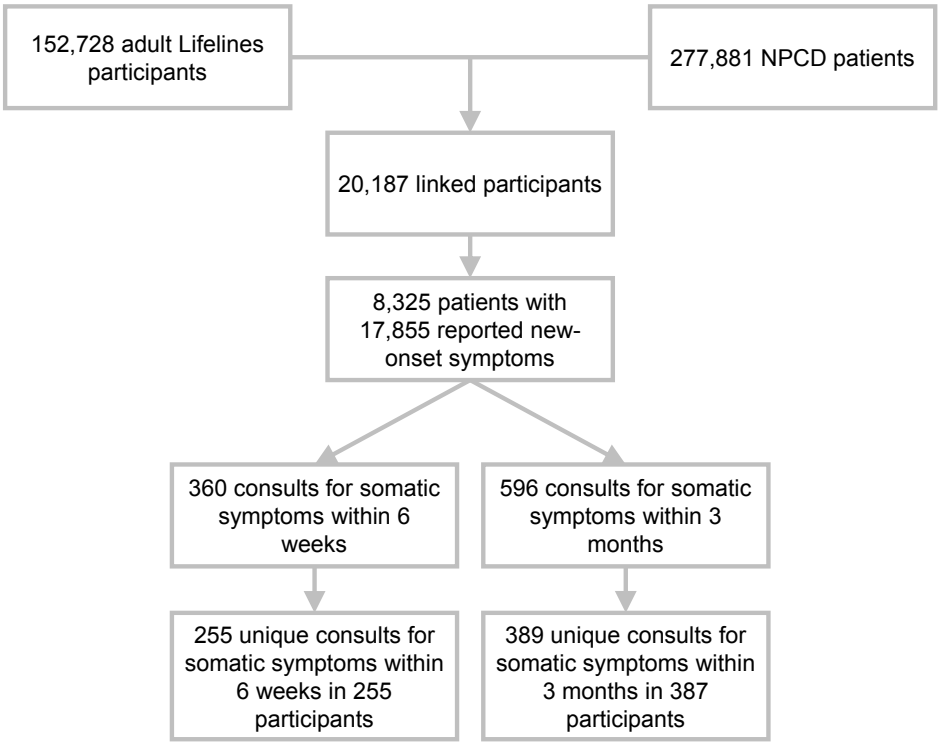


Figure C: Inclusion of participants and GP consults

PART 4

Diagnostics in primary care for common somatic symptoms

CHAPTER 8

Based on: Groeneveld, J.M., Ballering, A.V., van Boven, K., Akkermans, R.P., Olde Hartman, T.C., & Uijen, A.A. (2020). Sex differences in incidence of respiratory symptoms and management by general practitioners. *Family Practice*, *37*(5), 631-636.

Abstract

Background: Differences between women and men play an important role in lung physiology and epidemiology of respiratory diseases, but also in the health care processes.

Objective: To analyse sex differences in patients encountering their general practitioner (GP) with respiratory symptoms with regard to incidence, GP's management and final diagnoses.

Methods: Retrospective cohort study, using data of the Dutch Practice Based Research Network. All patients who encountered their GP from 01-07-2013 until 30-06-2018 with a new episode of care starting with a reason for encounter (RFE) in the respiratory category (R) of the ICPC-2 classification were included (n = 16,773). Multilevel logistic regression was used to analyse influence of patients' sex on management of GPs with adjustment for possible confounders.

Results: We found a significant higher incidence of respiratory symptoms in women than in men: 230/1000 patient years (95% CI 227-232) and 186/1000 patient years (95% CI 183-189), respectively. When presenting with cough, GPs are more likely to perform physical examination (OR 1.22; 95% CI 1.11-1.35) and diagnostic radiology (OR 1.25; 95%CI 1.08-1.44), but less likely to prescribe medication (OR 0.88; 95%CI 0.82-0.95) in men. When visiting the GP with dyspnoea, men more often undergo diagnostic imaging (OR 1.32; 95%CI 1.05-1.66) and are more often referred to a specialist (OR 1.35; 95% CI 1.13-1.62).

Conclusion: Women encounter their GP more frequently with respiratory symptoms than men and GPs perform more diagnostic investigations in men. We suggest more research in general practice focusing on sex differences and possible confounders.

Background

Men and women differ in their health and diseases. This variation is caused by biological characteristics such as anatomy or hormonal factors (ie, sex),[1] together with gender, meaning the different expected social roles, behaviors and cultural aspects related to being male or female.[2,3] The past years, growing scientific interest in the role of gender and sex on health care and specific diseases has developed.[4] Previous studies show several discrepancies in epidemiology and symptoms of conditions in males and females. However, these differences between sexes are not fully addressed by health care providers yet.[5,6]

Although health research about sex and gender has mainly focussed on cardiovascular conditions, recent evidence has shown that sex and gender influence lung physiology and respiratory diseases as well. This influence is seen throughout the whole lifespan. Starting in intra-uterine life, female fetuses show an earlier production of surfactant than males and have fewer number bronchi but these mature faster.[7,8] During childhood, boys have a higher prevalence of asthma than girls. This difference is probably due to relatively smaller airway diameters in males compared with females.[8] Asthma prevalence rises in females in puberty and decreases in men in puberty, reaching an equal prevalence for both sexes around the age of the menopause.[1] Also in chronic obstructive pulmonary disease (COPD), previously seen as a *'smoking men's disease'*, epidemiologic changes are noted. Incidence of COPD in females is rising, and evidence suggests that females are more susceptible to tobacco smoke than men.[7]

Sex and gender do not only influence epidemiology and pathogenesis of diseases, they also impact the actions of health care providers. Evidence shows that despite presenting with similar complaints in several conditions, women are less likely to undergo additional diagnostic investigations and are more often classified in a category of non-specific diagnoses than men.[9-11] Possible explanations for this phenomenon are that women tend to seek health care more often than men and have a less straightforward way of presenting their symptoms.[3,11]

Little to no research has yet been conducted into sex differences of specifically respiratory symptoms and subsequent management hereof by the general practitioner (GP). This is problematic, as sex is an inevitable determinant in research and in clinical practice. Additionally a lack of knowledge may lead to bias and suboptimal treatment when differences between males and females are not taken into account. In this study we aim to analyze the difference in incidence

of respiratory symptoms presenting to the GP between men and women, as these symptoms are a very common in primary care practice. Furthermore, we will analyze differences in the management and final diagnosis by the GP between men and women presenting with respiratory symptoms.

Methods

Design and data collection

This retrospective cohort study used electronic data from the Practice Based Research Network (PBRN) Family Medicine Network (FaMe-Net), a Dutch primary care research network from the Radboud University Medical Centre in Nijmegen.[12] Since 1971, all encounters between patients and GPs are registered in this network, which consists of seven family practices (24 GPs and approximately 32.000 registered patients). GPs routinely code episodes of care according to the International Classification of Primary Care (ICPC-2).[13,14] An episode of care is defined here as an individual health problem, that starts at the first encounter and is completed at the final encounter linked to that health problem. Furthermore, GPs register the patient's initial reason for encounter (RFE) for each episode, all performed interventions during the episode and the final GP's diagnosis.[15,16] Being the literal expression of the reason why patients encounter the GP, the RFE represents the demand of care for that person.[17] RFEs can be complaints and symptoms, but also a particular diagnosis or a request for an intervention, such as prescription of medication.

Population

We included patients of all ages who encountered their GP in the period 01-07-2013 until 30-06-2018 with a new episode of care starting with a RFE in the respiratory category (R) of the ICPC-2 (R-RFE). We excluded episodes of care that started solely with a request for intervention (R30-R69).

Measurements

We collected the following patient characteristics: sex, age at start of episode of care, GP practice and comorbidity. Relevant comorbidities were selected by their ICPC code: cardiovascular disease (K22, K72, K47-80, K82-84, K86-92, K99), COPD/chronic bronchitis (R79 and R95), asthma (R96) and presence of malignancies (A79, B72-74, D74-77, F74, H75, L71, N74, N76, R84, R85, T71-73, U75-77, U79, X75-77, X81, Y77-79). From encounters we collected the following information: the RFE, status of the visit (first encounter or subsequent encounter

within the episode), the type of encounter (consultation at the practice or at home, telephone or email consultation, both in daily practice as in evening or night shifts) and the final diagnosis of the episode of care.

Incidence of respiratory symptoms
We analysed the incidence (in patient years) of each RFE at exclusively the first visit of an episode, per sex and age category as used in previous research.[18,19] Patient years were extracted from the Electronic Medical Health Record TransHis, the information system of GPs participating in the PBRN FaMe-Net. When GPs coded more than one initial R-RFE at the start of an episode of care, we included every initial R-RFE in our analysis.

Management and final diagnosis of GPs
We focussed on the four R-RFEs with the highest incidence number, namely cough, dyspnoea, acute upper respiratory infection and throat symptoms, and analysed all interventions that were performed by the GP in the entire corresponding episodes of care. Interventions were grouped by their ICPC code: physical examination (-30 and -31), laboratory diagnostics (-33 and -34), diagnostic radiology/imaging (-41), medication prescription (-50), referral to other primary care provider (-66) or referral to specialized care/hospital (-67). Furthermore, we analysed for both sexes the final diagnosis of each episode of care started with the particular RFE. These diagnoses were coded by the ICPC-2 classification. The validity of registration of diagnoses is high, as participating GPs meet regularly to discuss registration and diagnostic criteria. Moreover, the electronic medical record system that was used, warns the GP in case of error or inconsistency in registration.

Data analysis
For data analysis we used tools provided in SPSS 25. We calculated incidence numbers and confidence intervals (CIs) using descriptive statistics. To investigate how patients' sex affects interventions delivered by GPs, we performed a multi-level analysis to determine the influence of variables on the presence or absence of an intervention. We corrected our findings for patients' sex, age, numbers of encounters in the episode and presence of comorbidities at start of the episode.

Results

We found 38,704 episodes of care starting with an R-RFE in 20,063 patients. We excluded 9063 episodes of care, because they started with a request for intervention as RFE (of which 66% was a request for influenza vaccination or a request for medication

prescription). Finally, we analysed 29,641 episodes of care in 16,773 patients. Baseline characteristics of the included patients are shown in ***Table 1***. Women encountered the GP more frequently with an R-RFE than men and women had more relevant comorbidities at the moment of encounter. We found no difference in the total number of encounters within an episode of care between male and female patients.

Table 1. Baseline characteristics of all patients encountering general practice with respiratory symptoms (01-07-2013 – 30-06-2018).

Characteristic	Men, no (%)	Women, no (%)	*p* value	Total
Patients	7594 (45%)	9179 (55%)		16,773
New episodes of care	13009 (44%)	16632 (56%)		29,641
Age at start of episode of care, mean (SD)	34 (27)	37 (25)		
Number of R-RFEs at start of episode of care, mean (SD)	1.15 (0.39)	1.17 (0.40)	<0.001*	
Number of encounters per episode, mean (SD)	1.61 (1.66)	1.60 (1.61)		
Comorbidity at start of episode of care	4269 (33%)	5712 (34%)	0.006*	9981
Cardiovascular disease	2999 (23%)	4002 (24%)	0.043*	7001
Asthma, COPD or chronic bronchitis	1751 (14%)	2202 (13%)	0,58	3953
Malignancy	623 (5%)	991 (6%)	<0.001*	1614

*$p<0.05$

The total incidence of R-RFEs is 208/1000 patient years (95% CI 206-210), with a significant difference in the incidence between men and women: 186/1000 patient years (95% CI 183-189) and 230/1000 patient years (95% CI 227-232), respectively. ***Figure 1*** shows the distribution of incidences for all R-RFEs per age category and sex. In the age category 0-4 years the incidence of R-RFEs is significantly higher in boys with an incidence of 537/1000 patient years (95% CI 525-550); for girls this incidence is 476/1000 patient years (95% CI 463-489). As age increases, the distribution of included R-RFEs per sex changes. The 10 most frequently coded R-RFEs for men and women are shown in ***Table 2***, with their corresponding incidences per 1000 patient years.

The four RFEs with highest incidences were cough (R05), dyspnoea (R02), throat symptoms (R21) and acute upper respiratory tract infection (R74). ***Table 3*** shows the interventions of GPs in episodes of care started with cough and dyspnoea. With a RFE cough, GPs more frequently perform a physical examination (OR 1.22; 95%CI 1.11-1.35), and diagnostic imaging (OR 1.25; 95%CI 1.08-1.44) in men and prescribe medication less often (OR 0.88; 95%CI 0.82-0.95) compared with women. When visiting the GP with RFE dyspnoea, the odds to receive diagnostic imaging in the episode of care is

1.32 times (95%CI 1.05-1.66) higher for male patients than for females. In addition, men presenting with dyspnoea are more often referred to a specialist in the episode of care (OR 1.35; 95%CI 1.13-1.62). In throat symptoms and acute upper respiratory infection no significant differences were found between male and female patients for all types of interventions (data not shown).

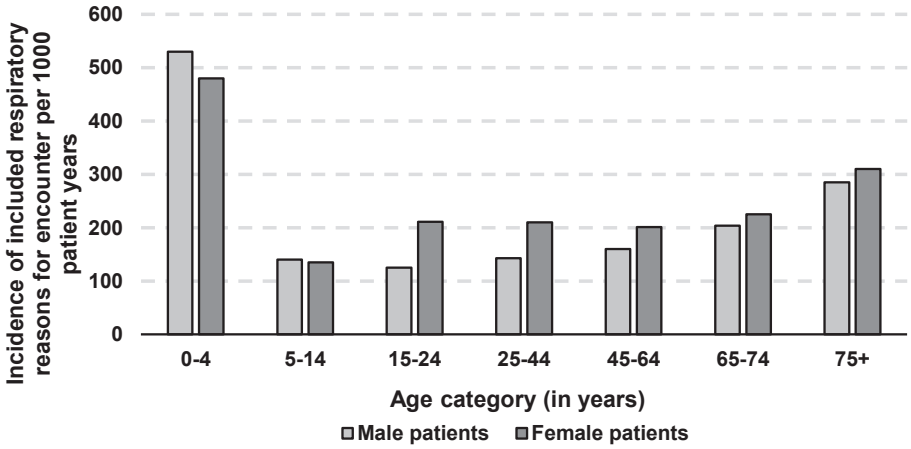


Figure 1. Total incidence number of all included respiratory reasons for encounter of general practice in the FaMe-Net database, divided per age category and sex (01-07-2013 – 30-06-2018).

Table 2. Incidence numbers of the ten most common respiratory reasons for encounter in general practice for both men and women (01-07-2013 – 30-06-2018).

Incidence per 1000 patient years (95% CI)			
Men		**Women**	
Total	186 (183-189)	Total	230 (227-232)
Cough (R05)	77.9 (76.0-79.8)	Cough (R05)	93.2 (91.1-95.2)
Throat symptoms (R21)	24.5 (23.4-25.6)	Throat symptoms (R21)	38.1 (36.8-39.4)
Dyspnoea (R02)	21.2 (20.2-22.3)	Dyspnoea (R02)	28.1 (26.9-29.2)
Acute upper respiratory tract infection (R74)	14.5 (13.6-15.3)	Acute upper respiratory tract infection (R74)	18.3 (17.4-19.3)
Sneezing (R07)	5.9 (5.3-6.4)	Sneezing (R07)	5.9 (5.4-6.5)
Epistaxis (R06)	5.0 (4.5-5.5)	Sinus symptoms (R09)	5.7 (5.2-6.2)
Nose symptoms (R08)	4.3 (3.8-4.7)	Nose symptoms(R08)	4.2 (3.8-4.7)
Breathing problem (R04)	2.3 (3.8-4.7)	Epistaxis (R06)	4.2 (3.7-4.6)
Wheezing (R03)	3.5 (3.1-3.9)	Sinusitis acute/chronic (R75)	4.1 (3.6-4.5)
Allergic rhinitis (R97)	3.4 (2.9-3.8)	Influenza (R80)	3.7 (3.2-4.1)

Table 3. Interventions performed by general practitioners in episodes of care starting with reasons for encounter 'cough' (R05) and 'dyspnoea' (R02) (01-07-2013 – 30-06-2018).

Intervention	Male patients (% of total)	Female patients (% of total)
Total episodes RFE cough	5903	7357
Physical examination	5129 (86.9%)	6192 (84.2%)
Laboratory diagnostics	992 (16.8%)	1335 (18.1%)
Diagnostic radiology/imaging	354 (6.0%)	349 (4.7%)
Medication prescription	2829 (47.9%)	3870 (52.6%)
Referral in first line	10 (0.2%)	23 (0.3%)
Referral to specialist/hospital	189 (3.2%)	204 (2.8%)
Total episodes RFE dyspnoea	1610	2217
Physical examination	1483 (92.1%)	2011 (90.7%)
Laboratory diagnostics	396 (24.6%)	619 (27.9%)
Diagnostic radiology/imaging	155 (9.6%)	175 (7.9%)
Medication prescription	789 (49.0%)	1016 (45.8%)
Referral in first line	56 (3.5%)	76 (3.4%)
Referral to specialist/hospital	289 (18.0%)	317 (14.3%)

*$p<0.05$ [a]Tested by multilevel logistic regression. [b]Adjusted for possible confounders: age, number of encounters in episode of care, cardiovascular comorbidity, asthma, COPD, chronic bronchitis, or known malignancy at moment of encounter.

The most frequent final diagnoses of the episodes of care starting with a cough and dyspnoea are presented in ***Table 4***. For both symptoms cough and dyspnoea, men are more often diagnosed with pneumonia. Women are more often diagnosed with 'sinusitis' when presenting with a cough, and with a final (symptom)diagnosis 'dyspnea' when presenting with dyspnea. For men presenting with cough, the diagnoses 'acute otitis media/myringitis', 'wheezing' and 'asthma' are more often assigned.

	Total episodes	OR *crude*[a] (95% CI)	OR *adjusted*[b] (95% CI)	*p* value, *adjusted*
	13,260			
	11321	1.25 (1.13-1.38)	1.22 (1.11-1.35)	<0.001*
	2327	0.92 (0.83 -1.01)	0.96 (0.87- 1.07)	0.47
	703	1.18 (1.02-1.35)	1.25 (1.08-1.44)	0.002*
	6699	0.83 (0.77-0.90)	0.88 (0.82 -0.95)	0.001*
	33	0.98 (0.80-1.20)	0.98 (0.80- 1.20)	0.83
	393	1.07(0.90-1.26)	1.05 (0.89-1.24)	0.54
	3827			
	3494	1.16 (0.93-1.46)	1.15 (0.92-1.44)	0.23
	1015	0.85 (0.73-0.98)	0.90 (0.77-1.05)	0.18
	330	1.20 (0.96-1.50)	1.32 (1.05-1.66)	0.016*
	1805	1.14 (0.99-1,30)	1.12 (0.98-1.29)	0.10
	132	1.01 (0.75-1.36)	1.04 (0.77-1.41)	0.79
	606	1.31 (1.10-1.57)	1.35 (1.13-1.62)	0.001*

Table 4. Final diagnoses of episodes of care starting with reasons for encounter 'cough' and 'dyspnoea' (01-07-2013 – 30-06-2018).

Final diagnosis (by ICPC-2)	Men, N (%)	Women, N (%)	*p* value
Episodes of care starting with cough (R05)			
R74 Acute upper respiratory infection	2340 (39.6%)	3001 (40.8%)	0.18
R05 Cough	1691 (28.7%)	2212 (30.1%)	0.075
R81 Pneumonia	526 (8.9%)	550 (7.5%)	0.003
R78 Acute bronchitis/bronchiolitis	411 (7.0%)	471 (6.4%)	0.20
R77 Acute laryngitis/tracheitis	186 (3.15%)	272 (3.7%)	0.087
R80 Influenza	166 (2.8%)	196 (2.7%)	0.60
H71 Acute otitis media/myringitis	92 (1.6%)	76 (1.0%)	0.007
R75 Sinusitis acute/chronic	46 (0.78%)	87 (1.2%)	0.021
R96 Asthma	64 (1.1%)	52 (0.71%)	0.020
R03 Wheezing	43 (0.73%)	26 (0.35%)	0.003
Episodes of care starting with dyspnoea (R02)			
R02 Dyspnoea	358 (22.2%)	613 (27.7%)	<0.001
R74 Acute upper respiratory infection	291 (18.1%)	416 (18.8%)	0.59
R81 Pneumonia	172 (10.7%)	181 (8.2%)	0.008
R78 Acute bronchitis/bronchiolitis	143 (8.9%)	164 (7.4%)	0.095
R96 Asthma	70 (4.4%)	94 (4.2%)	0.87
R77 Acute laryngitis/tracheitis	54 (3.4%)	71 (3.2%)	0.80
R98 Hyperventilation syndrome	47 (2.9%)	77 (3.5%)	0.34
R05 Cough	39 (2.4%)	49 (2.2%)	0.67
K77 Heart failure	40 (2.5%)	48 (2.2%)	0.52
R80 Influenza	30 (1.9%)	45 (2.0%)	0.71

Discussion

Summary

Female patients were found to have a significantly higher incidence of respiratory symptoms as RFE (230/1000 patient years) compared with male patients (186/1000 patient years). GPs perform different interventions in male and female patients presenting with the same RFE, especially in cough and dyspnoea. When presenting a cough, males are more likely to undergo physical examination and diagnostic radiology than females. Women, however, are more often prescribed medication for their cough than men. With regard to dyspnoea, males are more likely to undergo diagnostic radiology when encountering the GP and are more often referred to a

medical specialist during this episode of care than females. In the current research, these differences in GPs' actions are not explained by patient characteristics such as age, comorbidity or the number of encounters in the episode of care. Symptom-diagnoses such as 'dyspnoea' are more often assigned to women than to men, whose diagnoses are more clearly defined.

Strengths and limitations

A major strength of this study is the high validity and reliability of the FaMe-Net database, based on regular discussions between participating GPs to ensure conformity regarding registration. Additionally, we included many patients and encounters, allowing for greater statistical power. However, the retrospective study design has its limitations. First, differences in symptom-presentation between men and women may have had a substantial impact on GPs' decisions. In addition, the severity of the symptoms at the moment of presentation may vary between women and men. Health-care seeking behaviour may also influence decision making of GPs: women seek more care and at an earlier stage than men.[20] Other possible confounding factors are the personal conceptions and experiences of a GP, such as former patients with similar complaints and experienced benefits or harms of certain interventions. We could not take these possible confounding factors into account, due to the retrospective nature of the study. Furthermore, we could not correct our analyses for additional possible confounders including smoking status, socio-economic status (SES) and family history of respiratory disease. Especially the absence of information on patient's smoking status in the context of respiratory diseases is of importance, as this may have substantial effects on actions performed by the GP. Lastly, despite the high validity of ICPC-2 coding of RFEs and diagnoses in our study, potential misclassification of final disease or diagnoses may have occurred.

Comparison with previous literature

We found a total incidence of respiratory symptoms of 208/1000 patient years. This is in line with a Dutch study conducted in 2004,[19] which reports a total incidence of symptoms in the respiratory ICPC-category of 214/1000 patient years. Consistent with previous literature, we found that women seek more healthcare than men.[11,21] Both the absolute frequency of GP encounters, as the incidence of respiratory RFEs is higher in women than men. Possible explanations for this are differences in socialization patterns and cultural norms between men and women, allowing women to more easily seek healthcare.[22,23] At the ages of 0-4 years, incidence of respiratory symptoms is higher in male patients, but after puberty incidence is higher in female patients. These findings are in line with previous

research[19] which states that during childhood, male patients are more susceptible for respiratory disease due to anatomical development of the lungs, shifting to a higher prevalence of lung diseases in females from the age of puberty, possibly due to hormonal and physiological changes.[1,7]

Our results of differences in GP interventions match results of previous studies, in which women are less likely to receive more advanced diagnostic interventions in a large variety of diseases. Especially in cardiovascular disease, research shows that women are less likely than men to be diagnosed, treated and referred when experiencing chest pain.[3,10,24-26] These differences are seen prospectively in randomized trials and cohort studies, both in primary care and specialized care. Although in lesser amount, sex differences have been studied regarding respiratory symptoms or diseases.[9,27,28] However, no studies were performed in the Netherlands and the only study we found in primary care focusing on respiratory complaints has been performed in Spain with significantly fewer patients than this cohort study.[9]

A variety of factors may explain the differences in the nature and number of GP interventions. First, patients' biological factors might cause variations in type and severity of symptoms, which makes the GP decide upon different interventions. Also, additional patient characteristics including age, smoking status and family history are thought to be important factors for GPs to consider when deciding upon management of respiratory symptoms. However, after adjustment for age and the presence of relevant comorbidities, the influence of sex on GP's interventions remained present in our study. Second, differences in GPs' interventions might be explained by a differing likelihood for final diagnoses between men and women presenting with the same RFE. For example, men may receive the final diagnosis pneumonia more frequently than women. On the one hand, this could be due to a higher detection rate of pneumonia in men, as they receive more radiology than women. This situation may be self-sustaining: increased use of radiology in men leads to a higher detection rate of pneumonia in men compared with women. Consequently, the incidence of pneumonia in men increases, followed by an increased likelihood of GPs applying more radiology, as GPs base their actions on guidelines and personal experience. On the other hand, a truly higher incidence of pneumonia, irrespective of diagnostic procedures, might be present in men compared with women. The current research, however, cannot clarify which situation is more likely. Lastly, wishes and communication patterns of patients may contribute to differences in GP's interventions. The presentation of complaints of women is often considered to be more extensive and vague; men communicate more demanding and straightforward, possibly resulting in the demand for more thorough examinations or referral.[3]

Implications for further research and clinical practice

The findings of this study show differences in the incidence in respiratory symptoms and the interventions by GPs that follow these symptoms, between men and women. However, as we could not include some possible confounding variables, further research is necessary to assess to which extent a sex bias contributes to these differences. We suggest future research that assesses possible factors that influence GPs' interventions. Such research could include patient's gender, smoking habits, SES and family history of respiratory disease, but also communication aspects and the sex of the GP. Furthermore, qualitative research in patients with respiratory symptoms who are considering seeking help from a GP and in patients who actually did seek help, may provide information on patients' expectations of and experiences with their GP and whether these expectations and experiences differ between women and men. Additionally, in clinical practice, more awareness of the influence of patients' sex could be raised when GPs are taught to reflect on their actions focusing on possible sex- and communication-related aspects. Results of studies focusing on sex differences could be incorporated in training programmes for GPs and in the curriculum of medical students and GP trainees. In conclusion, our study suggests the need for more awareness of sex differences in primary care itself, as well as in primary care research.

References

1. Fuseini H, Newcomb DC. Mechanisms driving gender differences in asthma. *Current Allergy and Asthma Reports*. 2017;17(3):19. doi: 10.1007/s11882-017-0686-1.
2. Lagro-Janssen T, Lo Fo Wong S, van den Muijsenbergh M. The importance of gender in health problems. *European Journal of General Practice*. 2008;14:33-37. doi: 10.1080/13814780802436127.
3. Hamberg K. Gender bias in medicine. *Womens Health*. 2008;4(3):237-243. doi: 10.2217/17455057.4.3.237.
4. Jahn I, Börnhorst C, Günther F, Brand T. Examples of sex/gender sensitivity in epidemiological research: Results of an evaluation of original articles published in JECH 2006–2014. *Health Research Policy and Systems*. 2017;15(1):11. doi: 10.1186/s12961-017-0174-z.
5. Nederlands Huisartsgenootschap. Totaalbestand nationale onderzoeksagenda huisartsgeneeskunde -alle aangeleverde kennislacunes en kennisvragen. *Nederlands Huisartsgenootschap (NHG)*. 2017.
6. Nederlands Huisartsen Genootschap in samenwerking met de werkgroep Research van het Interfacultair Overleg Huisartsgeneeskunde, (IOH-R). Nationale onderzoeksagenda huisartsgeneeskunde. *Nederlands Huisartsgenootschap (NHG)*. 2018.
7. Townsend EA, Miller VM, Prakash YS. Sex differences and sex steroids in lung health and disease. *Endocrine reviews*. 2012;33(1):1-47. doi: 10.1210/er.2010-0031.
8. Carey MA, Card JW, Voltz JW, et al. It's all about sex: Gender, lung development and lung disease. *Trends in endocrinology and metabolism*. 2007;18(8):308-313. doi: 10.1016/j.tem.2007.08.003.
9. Ruiz-Cantero MT, Ronda E, Álvarez-Dardet C. The importance of study design strategies in gender bias research: The case of respiratory disease management in primary care. *Journal of epidemiology and community health*. 2007;61(Suppl 2):ii11-ii16. doi: 10.1136/jech.2007.060301.
10. Arber S, McKinlay J, Adams A, Marceau L, Link C, O'Donnell A. Patient characteristics and inequalities in doctors' diagnostic and management strategies relating to CHD: A video-simulation experiment. *Social Science & Medicine*. 2006;62(1):103-115. doi: 10.1016/j.socscimed.2005.05.028.
11. Loikas D, Karlsson L, von Euler M, Hallgren K, Schenck-Gustafsson K, Bastholm Rahmner P. Does patient's sex influence treatment in primary care? experiences and expressed knowledge among physicians--a qualitative study. *BMC family practice*. 2015;16:137. doi: 10.1186/s12875-015-0351-5.
12. Uijen AA, Bor H, van Boven K. FaMe-net: Twee oude registratienetwerken in een nieuw jasje. *Tijdschrift voor Gezondheidswetenschappen*. 2015;93(8):286-287.
13. Hofmans-Okkes I, Lamberts H. The international classification of primary care (ICPC): New applications in research and computer-based patient records in family practice. *Family practice*. 1996;13(3):294-302. doi: 10.1093/fampra/13.3.294.
14. Soler JK, Okkes I. Reasons for encounter and symptom diagnoses: A superior description of patients' problems in contrast to medically unexplained symptoms (MUS). *Family practice*. 2012;29(3):272-282. doi: 10.1093/fampra/cmr101.
15. van den Broek J, van Boven K, Bor H, Uijen AA. Change in frequency of patient requests for diagnostic screening and interventions during primary care encounters from 1985 to 2014. *Family practice*. 2018;35(6):724-730. doi: 10.1093/fampra/cmy031.
16. van Boven K, Uijen AA, van de Wiel N, Oskam SK, Schers HJ, Assendelft WJJ. The diagnostic value of the patient's reason for encounter for diagnosing cancer in primary care. *Journal of the American Board of Family Medicine*. 2017;30(6):806-812. doi: 10.3122/jabfm.2017.06.170076 [doi].

17. Bentzen N. Family medicine research: Implications for wonca. *Annals of Family Medicine*. 2004;2(Suppl 2):S45-S49. doi: 10.1370/afm.190.
18. Hak E, Rovers M, Kuyvenhoven M, Schellevis F, Verheij T. Incidence of GP-diagnosed respiratory tract infections according to age, gender and high-risk co-morbidity: The second dutch national survey of general practice. *Family practice*. 2006;23(3):291-294. doi: 10.1093/fampra/cmi121.
19. van der Linden MW, Westert GP, de Bakker DH, Schellevis FG. Tweede nationale studie naar ziekten en verrichtingen in de huisartspraktijk. klachten en aandoeningen in de bevolking en in de huisartspraktijk. *NIVEL/RIVM*. 2004.
20. Thompson AE, Anisimowicz Y, Miedema B, Hogg W, Wodchis WP, Aubrey-Bassler K. The influence of gender and other patient characteristics on health care-seeking behaviour: A QUALICOPC study. *BMC family practice*. 2016;17(1):38.
21. Osika Friberg I, Krantz G, Määttä S, Järbrink K. Sex differences in health care consumption in sweden: A register-based cross-sectional study. *Scandinavian Journal of Public Health*. 2016;44(3):264-273. doi: 10.1177/1403494815618843.
22. Schenck-Gustafsson KDP, Pfaff D. *Handbook of clinical gender medicine*. Basel: Karger; 2012.
23. Galdas PM, Cheater F, Marshall P. Men and health help-seeking behaviour: Literature review. *Journal of Advanced Nursing*. 2005;49(6):616-623. doi: 10.1111/j.1365-2648.2004.03331.x.
24. Nyberg F, Osika I, Evengård B. "The laundry bag project"– unequal distribution of dermatological healthcare resources for male and female psoriatic patients in sweden. *International journal of dermatology*. 2008;47(2):144-149. doi: 10.1111/j.1365-4632.2008.03485.x.
25. Clerc Liaudat C, Vaucher P, De Francesco T, et al. Sex/gender bias in the management of chest pain in ambulatory care. *Womens Health*. 2018;14:1745506518805641.
26. Pagidipati NJ, Coles A, Hemal K, et al. Sex differences in management and outcomes of patients with stable symptoms suggestive of coronary artery disease: Insights from the PROMISE trial. *American Heart Journal*. 2019;208:28-36. doi: 10.1016/j.ahj.2018.11.002.
27. Delgado A, Saletti-Cuesta L, López-Fernández LA, Gil-Garrido N, de Dios Luna del Castillo J. Gender inequalities in COPD decision-making in primary care. *Respiratory medicine*. 2016;114:91-96. doi: 10.1016/j.rmed.2016.03.017.
28. Roberts NJ, Patel IS, Partridge MR. The diagnosis of COPD in primary care; gender differences and the role of spirometry. *Respiratory medicine*. 2016;111:60-63. doi: 10.1016/j.rmed.2015.12.008.

CHAPTER 9

Ballering, A.V., Muijres, D., Uijen, A.A., Rosmalen, J.G.M., & olde Hartman, T.C. (2021). Sex differences in the trajectories to diagnosis of patients presenting with common somatic symptoms in primary care: an observational cohort study. *Journal of Psychosomatic Research*, *149*, 110589.

Abstract

Objective: Little insight exists into sex differences in diagnostic trajectories for common somatic symptoms. This study aims to quantify sex differences in the provided primary care diagnostic interventions for common somatic symptoms, as well as the consequences hereof for final diagnoses.

Methods: In this observational cohort study, we used data from the Dutch Family Medicine Network (N=34,268 episodes of care related to common somatic symptoms; 61,4% female). The association between patients' sex on the one hand, and diagnostic interventions and disease diagnoses on the other hand, were assessed using multilevel multiple logistic regression analyses. Structural equation modelling was used to estimate a mediation model with multiple parallel mediators to assess whether the fewer disease diagnoses given to female patients were mediated by the fewer diagnostic interventions female patients receive, compared to male patients.

Results: Women received fewer physical examinations (OR=0.84, 95%CI=0.79-0.89), diagnostic imaging (OR=0.92, 95%CI=0.84-0.99) and specialist referrals (OR=0.85, 95%CI=0.79-0.91) than men, but more laboratory diagnostics (OR=1.27, 95%CI=1.19-1.35). Women received disease diagnoses less often than men for their common somatic symptoms (OR=0.94, 95%CI=0.89-0.98). Mediation analysis showed that the fewer disease diagnosis in female patients were mediated by the fewer diagnostic interventions conducted in women compared to men.

Conclusion: This study shows that sex inequalities are present in primary care diagnostic trajectories of patients with common somatic symptoms and that these lead to unequal health outcomes in terms of diagnoses between women and men. FPs have to be aware of these inequalities to ensure equal high-quality care for all patients.

Introduction

Health outcomes are closely related to patients' trajectories to diagnosis (1). A diagnostic trajectory comprises everything a family physician (FP) does, including diagnostic interventions, to obtain a diagnosis for a patient's complaint. Yet, patients' characteristics, including their sex, may influence the FPs perception of symptoms and consequently, patients' diagnostic trajectories.[1]

Diagnostic trajectories for multiple diseases differ between women and men. A recent study shows that women receive fewer physical examinations, diagnostic imaging and specialist referrals when they present with cough and/or dyspnoea in primary care than men.[2] Additionally, studies show that sex is associated with different diagnostic trajectories in coronary heart disease (CHD). Women presenting with symptoms suggestive of CHD in primary care are less likely to receive physical examinations.[3,4] A similar sex difference is observed in colorectal cancer, as studies show women have higher mortality rates and are less often screened than men.[5] However, a recent study shows that men are less likely to receive an early diagnosis of dementia than women.[6]

Although sex differences in the prevalence and longevity of common somatic symptoms are recognized,[7,8] little insight exists into sex differences in the primary care diagnostic trajectories related to these symptoms. Most research into diagnostic trajectories focuses on previously diagnosed disease[9,10] and only the aforementioned study on respiratory symptoms studied diagnostic trajectories from symptom presentation to final diagnosis in primary care.[2] Additionally, previous research into diagnostic trajectories is based on self-reported outcomes, which are prone to recall bias.[11,12]

It is thought that sex differences exist in whether patients receive a disease diagnosis,[13,14] with women's symptoms remaining more often unexplained than men's. However, it remains unknown whether the primary care diagnostic trajectory associates with symptoms that remain unexplained. It is pivotal to understand sex differences in diagnostic trajectories of somatic symptoms as these may result in sex inequalities in healthcare and sex-skewed morbidity or mortality rates.[5,15]

Therefore, this study aims to quantify sex differences in primary care diagnostic trajectories of patients with common somatic symptoms and to assess whether potential sex differences in these trajectories are associated with whether the somatic symptoms are ultimately attributable to a disease (i.e. disease diagnosis) or continue to be symptoms that cannot be (fully) explained by an underlying disease

or bodily abnormality (i.e. symptom diagnosis).[16] Firstly, we will study the presence of differences between female and male patients in provided diagnostic interventions when a patient presents with somatic symptoms in primary care. Secondly, we will assess whether sex differences exist in whether the patient receives a disease diagnosis when presenting with somatic symptoms. Lastly, we will study whether potential differences in women's and men's diagnostic trajectories are associated with patients' final diagnoses.

Methods

Study design and data collection

In this observational cohort study we used data from the Practice Based Research Network (PBRN) Family Medicine Network (FaMe-Net), which includes approximately 32,000 patients and 26 FPs in seven different primary care practices throughout the Netherlands. Studies involving FaMe-Net data are exempted from ethical review by the CCMO (Dutch Central Committee on research involving human subjects). Patients are extensively informed about the inclusion of their health-related information in FaMe-Net, and are offered the opportunity to opt out of FaMe-Net. FaMe-net is embedded in the regular Dutch primary care system. It is the world's oldest PRBN and since its inception patients' morbidity is systematically registered within an episode of care (EoC) structure.[17] An EoC is defined as a patient's health problem from the first encounter until the last encounter related to that specific health problem. Within each encounter of an EoC, the FP routinely and systematically codes the patient's reason for encounter (RFE), diagnosis and interventions (including physical examinations, diagnostic tests and referrals to specialists) according to the International Classification of Primary Care (ICPC-2). An EoC may start with more than one RFE, but one final diagnosis is ultimately linked to all encounters within the EoC. The registered diagnosis of an EoC can be modified anytime when new insights regarding the patient's RFE arise. The RFE should be acknowledged by the patient as a correct description of their demand of care. The validity of data registration is high, as participating FPs structurally meet to discuss diagnostic criteria to minimize bias during registration. Moreover, the automated FP information system recognizes inconsistencies in registration.

We selected EoC that started on January 1st, 2014 until December 31st, 2018 with at least one common somatic symptom. Contacts within EoC that continued hereafter were excluded. Face-to-face encounters, telephone and digital consultations were included within an EoC. The relative distribution of the type of contact was similar

among female and male EoC. We included fifteen RFEs related to twelve symptoms: headache (ICPC-N01), dizziness (ICPC-N17), heart pain (ICPC-K01), (lower) back pain (ICPC-L02 and ICPC-L03), nausea (ICPC-D09), muscle pain (ICPC-L18), shortness of breath/dyspnoea (ICPC-R02), chills (ICPC-A02), tingling of fingers, feet and/or toes (ICPC-N05), swallowing/throat problems (ICPC-D21 and ICPC-R21), weakness or general tiredness (ICPC-A04), and arm or leg symptoms (ICPC-L09 and ICPC-L14). These symptoms reflect the contents of the Symptom CheckList-90 Somatization subscale (SCL-90 SOM),[18] are common,[19] and often remain unexplained.[1] EoCs starting with the same RFE on the same date within the same patient were excluded (N=106). We analysed the EoC that started with more than one RFE related to somatic symptoms (N=1,605; 4.7%) as an EoC of the first-mentioned RFE.

Statistical Analyses

To assess sex differences in whether a physical examination (ICPC-30 and ICPC-31) was conducted, and whether laboratory diagnostic interventions, including microbiological/immunological testing and blood tests (ICPC-33 and ICPC-34, respectively), diagnostic imaging (ICPC-41) and specialist referrals (ICPC-67) were requested when patients presented common somatic symptoms, we conducted multilevel multiple logistic regression analyses. Patients' sex, patients' age at time of diagnosis, the number of contacts between patients and FPs during an EoC, the type of RFE, the type of consults (face-to-face or by phone/electronic) and the presence of comorbidities at the start of an EoC (***Appendix A***) were included as independent variables; the diagnostic interventions were included as dependent variable. As EoCs are nested at the individual level, analyses were clustered at this level. To exclude that the association between sex and laboratory diagnostics was explained by women receiving laboratory diagnostics to confirm a urinary tract infection when presenting (lower) back pain, we conducted a sensitivity analyses in which EoC starting with (lower) back pain were excluded. We included sex-by-RFE interaction terms to assess whether the association between the type of RFE and interventions differed between female and male patients.

Similar analyses, with disease diagnosis as the outcome, were conducted to assess whether the presence of disease diagnoses in an EoC that started with somatic symptoms differed between women and men. A disease diagnosis was operationalized as ICPC≥70, including psychiatric ICPC codes. This means that symptoms, followed over time, evolved in a diagnosed disease. In contrast, a symptom diagnosis was operationalized as ICPC≤30, in which symptoms followed over time continued to be symptoms as relevant diagnostic criteria were not met. For example, a symptom diagnosis is registered for symptoms if during the whole year no medical diagnosis (i.e. an ICPC≥70) has been registered as explanation for the symptom.

To assess whether the diagnostic interventions mediated the association between sex and disease diagnosis, we used structural equation modelling to estimate a mediation model with multiple parallel mediators. We used weighted least squares-estimation with 1000 bootstraps. Because all involved variables were binary, we used a probit link-function to estimate all regression coefficients and 95%CIs for direct and indirect paths between sex and final diagnosis. To facilitate easier interpretation of the probit slopes, we converted the predicted probabilities of the probit model to odds and computed the concomitant odds ratio.[20,21]

To test whether continuous covariates included in the logistic regression analyses fulfilled the linearity assumption of multiple logistic regression we divided the covariates into categories, and assessed whether the estimates changed monotonically. We found no indication for multicollinearity as the VIF was <5 in all analyses.[22] The statistical analyses and additional descriptive analyses were conducted in IBM SPSS Statistics v. 25, except for the mediation analysis, which was performed in R version 3.5.1, package '*Lavaan*' version 0.6-5.[23] We maintained a two-sided alpha-value of $p<0.004$ corrected for multiple testing.

Results

Study population and incidence rates of common somatic symptoms

We identified 34,268 unique EoC that started with 46,898 RFEs. In total, 9690 (28.3%) EoC started with more than one RFE. An overview of the study population is given in ***Table 1***.

Table 1. Overview of the study population

Characteristic	Women	Men	*p*-value
Patients, N (%)	10,541 (57.1)	7,915 (42.9)	
Patient years[a], N (%)	78,954 (52.3)	75,734 (47.7)	
New EoC, N (%)	21,031 (61.4)	13,237 (38.6)	
Age in years, Mean (SD)	43.4 (23.1)	42.3 (23.9)	**<0.001[b]**
Number of encounters per EoC, Median (IQR)	1.0 (1.0, 2.0)	1.0 (1.0, 2.0)	0.053[c]
EoC with comorbidities at the start, N (%)	8,282 (39.4)	5,295 (43.7)	0.271[d]
Cardiovascular disease	5,741 (27.3)	3,827 (28.9)	**0.001[d]**
Asthma and/or COPD	3,106 (14.8)	1,998 (15.1)	0.427[d]
Malignancies	1,684 (8.0)	979 (7.4)	0.038[d]

[a]The cumulative years the included patients were at risk for an incident common somatic symptom.
[b]Independent T-test; [c]Mann-Whitney *U* test; [d]Chi-Square test; bold indicates statistical significance

Figure 1 shows that heart pain and chills were equally presented by women and men; other symptoms were more frequently presented by women than men. The most prevalent symptoms in both sexes were weakness/general tiredness, (lower) back pain, swallowing/throat symptoms, and arm or leg symptoms. In total, 21,031 (61.4%) female EoC and 13,327 (38.6%) male EoC were found, translating into a total of 266.4 (95%CI=263.3-269.5) and 174.8 (95%CI=172.1-177.5) EoC related to somatic symptoms per 1000 patient years in women and men, respectively (***Appendix B***).

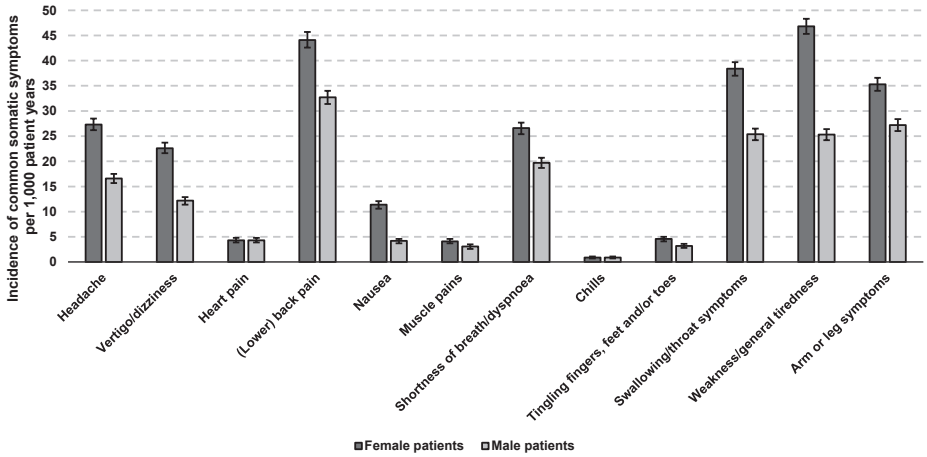


Figure 1. Incidence of common somatic symptom, stratified by sex.

Sex differences in diagnostic interventions

Table 2 shows that women have 0.84 (95%CI=0.79-0.89) times the odds of men to receive a physical examination during their EoC, adjusted for age, number of contacts within an EoC, type of consult and RFE, and comorbidities. Also, female patients received fewer diagnostic imaging and specialist referrals than their male counterparts. In contrast, we found that women were more likely to receive laboratory diagnostics than men (OR=1.27, 95%CI=1.19-1.35). We found an unadjusted OR of 1.28 (95%CI=1.21-1.35) and an adjusted OR of 1.20 (95%CI=1.12-1.28) in sensitivity analyses that excluded EoC related to (lower) back pain.

In models in which we included sex-by-RFE interaction terms we found that the associations between the type of RFE and diagnostic imaging or a specialist referral did not differ between women and men. However, a significant interaction term (OR=0.33, 95%CI=0.13-0.81) showed that women were less likely to receive a physical examination when presenting chills. Moreover, significant interaction terms showed

that women were more likely to receive laboratory diagnostics if they presented with vertigo/dizziness, (lower) back pain and weakness/general tiredness than men (OR=1.91, 95%CI=1.43-2.56; OR=1.69, 95%CI=1.26-2.26; OR=1.46, 95%CI=1.14-1.88, respectively). Analyses stratified per RFE are shown in ***Appendix C***.

Table 2. Associations between sex and the interventions conducted during an EoC.

Intervention	**Number of EoC (%)**		**OR (95%CI)[a]**	
	Female EoC	**Male EoC**	**Unadjusted**	**Adjusted[b]**
Physical examination	16,052 (78.5)	10,876 (82.1)	0.79 (0.75-0.84)	0.84 (0.79-0.89)
Laboratory diagnostics	5,649 (26.9)	2,875 (21.7)	1.33 (1.26-1.40)	1.27 (1.19-1.35)
Diagnostic imaging	1,264 (6.0)	951 (7.2)	0.83 (0.76-0.90)	0.92 (0.84-0.99)
Specialist referrals	2,160 (10.3)	1,611 (12.2)	0.83 (0.77-0.89)	0.85 (0.79-0.91)

[a]Odds ratios reflect women compared to men. [b]Adjusted for the patients' age at time of diagnosis, number of contacts in an EoC, presence of comorbidities at the start of an EoC, type of consult and the type of RFE.

Sex differences in final diagnoses

Table 3 shows that women have 0.94 times the odds (95%CI=0.89-0.98) of men to receive a disease diagnosis, when they present with somatic symptoms in primary care, adjusted for their age, number of contacts within an EoC, type of consult and RFE and comorbidities. In heart pain we found the largest sex difference in receiving a disease diagnosis as final diagnosis: women had 0.56 times the odds (95%CI=0.37-0.84) compared to men to receive a disease diagnoses. ***Appendix D*** provides the sex-stratifed final diagnoses most commonly given for somatic symptoms.

Mediation effects of interventions on the association between sex and disease diagnoses

Table 4 shows that the association between sex and disease diagnoses (i.e. the direct effect) is no longer statistically significant (OR=1.02; 95%CI=0.99-1.06), whereas the total effect was statistically significant (OR=0.94; 95%CI=0.93-0.98). Diagnostic interventions, except for diagnostic imaging, at least partly explain the association between sex and final diagnosis. ***Appendix E*** shows the probit regression coefficients.

Table 3. Associations between sex and receiving a disease diagnosis when presenting common somatic symptoms in primary care, stratified per RFE.

	EoC with disease diagnoses, N (%)		OR (95%CI)[a]	
RFE	**Female EoC**	**Male EoC**	**Unadjusted**	**Adjusted[b]**
Headache	858 (39.9)	506 (40.2)	0.99 (0.86-1.14)	0.99 (0.85-1.14)
Dizziness	629 (35.3)	365 (39.3)	0.84 (0.72-0.99)	0.86 (0.73-1.01)
Heart pain	60 (17.8)	93 (28.8)	0.54 (0.37-0.77)	0.56 (0.37-0.84)
(Lower) back pain	596 (17.1)	360 (14.5)	1.21 (1.05-1.40)	1.20 (1.03-1.39)
Nausea	295 (32.7)	114 (35.5)	0.88 (0.68-1.15)	0.87 (0.66-1.15)
Muscle pain	50 (15.5)	51 (22.0)	0.65 (0.42-1.00)	0.65 (0.41-1.02)
Shortness of breath	1,240 (59.0)	933 (62.5)	0.87 (0.76-0.99)	0.92 (0.80-1.06)
Chills	35 (52.2)	40 (61.5)	0.68 (0.34-1.37)	0.71 (0.32-1.60)
Tingling fingers, feet and/or toes	149 (40.8)	101 (41.4)	0.98 (0.70-1.36)	0.96 (0.68-1.35)
Swallowing/throat symptoms	1,821 (60.0)	1,108 (57.5)	1.11 (0.99-1.25)	1.21 (1.06-1.36)
Weakness/general tiredness	698 (18.9)	440 (23.0)	0.78 (0.68-0.89)	0.77 (0.67-0.89)
Arm and/or leg symptoms	1,039 (37.3)	813 (39.5)	0.91 (0.81-1.03)	0.89 (0.79-1.00)
Total	7,470 (35,5)	4,924 (37.2)	0.93 (0.89-0.97)	0.94 (0.89-0.98)[c]

[a]Odds ratios reflect women receiving a disease diagnosis (either psychiatric or somatic) compared to men. [b]Adjusted for the patients' age at time of diagnosis, number of contacts in an EoC, presence of comorbidities at the start of an EoC and type of consult. [c]Adjusted for aforementioned factors and type of RFE as well.

Table 4. Parallel mediation effects of interventions on the association between female sex and final diagnosis for all RFEs.

		OR (95%CI)[a,b]	
Intervention	**Indirect effect**	**Direct effect**	**Total effect**
Physical examination	0.96 (0.94-0.97)		
Imaging	1.01 (1.00-1.01)		
Laboratory diagnostics	0.96 (0.95-0.97)		
Specialist referral	0.98 (0.98-0.99)		
Total	0.91 (0.90-0.93)	1.02 (0.99-1.06)	0.94 (0.93-0.98)

[a]Odds ratios reflect women compared to men. [b]Adjusted for patient's age at time of diagnosis, number of contacts in an EoC, presence of comorbidities at the start of an EoC and the type of RFE.

Discussion

This study identified sex differences in the diagnostic trajectories of patients presenting with somatic symptoms in primary care. We found that women present themselves more often with somatic symptoms than men, and that men are more likely to receive a physical examination, diagnostic imaging and specialist referrals than women, whereas women received more laboratory diagnostics. This is the first study that shows that the fewer disease diagnoses given to women are mediated by the fewer diagnostic interventions that are performed in women.

Strengths and limitations

Our study had several strengths. First, the FaMe-Net database holds rich, valid and accurate primary care data of a large patient cohort. Participating FPs frequently discuss coding and the inherent warning system in the FP information system resulted in no indication of missing data in our cohort. Additionally, we used the RFE to define symptoms, which is an accurate measure, as the RFE is not interpreted by the FP.[24]

We acknowledge several limitations of our study as well. First, the patients' full medical history was unknown to the researchers. Second, the adjustment for comorbidities included cardiovascular diseases, respiratory diseases and malignancies, which is non-exhaustive. Both the patient's full medical history and comorbidities could have affected FPs' clinical decision making process and final diagnosis. Further research should investigate whether FPs value medical history and comorbidities similarly in female and male patients. Also, we could not include possible effect modifiers in our analyses, for example the patient's socioeconomic status or ethnicity, and the FP's working experience and sex,[25] as these data were unavailable.

Notably, especially the FP-patient sex concordance is important in patients' diagnostic trajectories. For example, female physicians are more inclined to deliver female preventive procedures, such as a Pap-smear.[25,26] Patients, irrespective their sex, are also more assertive and demanding towards female FPs than towards male FPs. Therefore, patients may request more interventions from female FPs.[27] Thus, if more female than male FPs are included in our study, the results may be skewed towards more diagnostic interventions, especially in female patients.

Comparison with existing literature

In line with previous studies, we found women to present more somatic symptoms in primary care than men.[28,29] We also found women to receive fewer diagnostic interventions, except for laboratory diagnostics, when presenting common somatic

symptoms in primary care. Our results are in line with a recent study that assessed patients' diagnostic trajectories in primary care for respiratory complaints: this study found that women receive fewer physical examinations, diagnostic imaging and specialist referrals than men when presenting with cough and/or dyspnoea.[2] Additionally, previous studies assessing the management of CHD found similar sex differences in the conduct of physical examinations and specialist referrals.[4,30,31] One study, in contrast, found a higher diagnostic test ordering in primary care in men for CHD-related symptoms.[3] However, this study did not distinguish between diagnostic imaging and laboratory diagnostics.

Our finding that women are more often diagnosed with a symptom diagnosis than men is in line with earlier literature[32]: a previous study found that women have 0.82 (95%CI=0.74-0.88) times the odds of men to receive a disease diagnosis for their somatic symptoms.[1] The difference in strength of the association between female sex and a disease diagnosis compared to the current study may be partly due to the different symptoms that were assessed in the previous study. The previous study included, for example, constipation and sleep disturbances, whereas the current study did not include these symptoms. Similarly, a recent study found an incidence rate ratio of 0.70 (95%CI=0.54-0.91) for receiving a disease diagnosis for somatic symptoms in women compared to men in a Latino and Asian American patient population.[33] Possibly, ethnic diversity interacts with sex, resulting in an excess of symptom diagnoses in women.

It is argued that many factors interact and form a complex interplay that could partially explain the sex differences we found in interventions and final diagnoses. First, it is known that women report more numerous and varied somatic symptoms,[7,34] possibly due to biological sex differences in the central processing of sensory information.[35]

Second, women visit their FP more often and earlier in their symptom trajectory than men.[36] The heightened bodily vigilance of women and their familiarization with primary care is thought to lower the threshold of seeking care for somatic symptoms.[37,38] In contrast, masculine gender stereotypes, for example stoicism about symptom reporting,[28,39] may delay male healthcare-seeking behaviour. Due to the relatively early female healthcare-seeking behaviour, women present with less typical symptoms of disease and consequently receive fewer specialist referrals and diagnoses.

Third, sex differences in communication styles could contribute to the sex difference in performed diagnostic interventions: women's communication is more subjective, talkative and polite, whereas men's communication is thought to be more demanding, rational and straightforward.[40]

Fourth, we found that women receive more laboratory diagnostics. The higher likelihood of requesting laboratory diagnostics for female patients could be an attempt to reassure the patient, as uncertainty about symptoms is more often reported in women.[3,41,42] However, FPs also request diagnostic tests to mitigate their own diagnostic uncertainty,[43,44] which is thought to be higher in female patients. Notably, sex differences in diagnostic uncertainty are mainly investigated for CHD-related symptoms[3,45]; to our knowledge no studies investigated sex differences in diagnostic uncertainty in common somatic symptoms. Additionally, laboratory diagnostics are an easily conducted and accessible intervention to rule out underlying conditions related to common somatic symptoms that are more frequent in women, for example urinary tract infections and anaemia.

Last, FPs' prejudices could also promote sex differences in diagnostic trajectories. For example, it is argued that physicians more readily assume somatization in women and offer fewer interventions to women.[46] FPs' state that they feel that their unconscious presumptions about gender and sex affect their clinical decision making process in a way that may disadvantage one sex over the other.[37]

Conclusions and implications for research and practice

This study shows that sex inequalities exist in the diagnostic trajectories of patients presenting somatic symptoms in primary care. However, it remains unknown whether these inequalities are justified. Although we used men as reference category in our analyses, male diagnostic trajectories are not necessarily the golden standard: possibly males are overmedicalized in comparison to females. The sex differences in diagnostic interventions may also be due to a differing probability for final diagnoses between women and men. Men receive disease diagnoses more frequently than women. On the one hand, this could be attributable to a higher detection rate of diseases in men as they receive more diagnostic interventions. Ultimately, this may be a vicious circle: increased use of diagnostic interventions in male patients leads to a higher detection rate of diseases in men compared to women. Consequently, the male rate of disease diagnoses increases, followed by an increased probability of FPs ordering more diagnostic interventions, as FPs base their actions on guidelines and personal experience. On the other hand, men may have a truly higher *a priori* chance of underlying disease than women when consulting the FP for somatic symptoms, which justifies the sex difference in diagnostic interventions. Therefore, further research could focus on whether women and men equally benefit from diagnostic interventions. Additionally, an experimental vignette-based study in which the vignette differs in patient's sex that asks FPs to suggest diagnostic interventions for the patient with a common somatic symptom may provide insights into sex differences, and reasons behind these, in diagnostic interventions suggested by FPs.[47]

Clinically this study suggest that the fewer diagnostic interventions in women may result in fewer disease diagnoses. Thus, if additional research shows that men and women equally benefit from diagnostic interventions in terms of receiving a disease diagnosis, FPs should be aware of the negative impact of the underuse of diagnostic interventions in female patients.

References

1. Verhaak PF, Meijer SA, Visser AP, Wolters G. Persistent presentation of medically unexplained symptoms in general practice. *Family practice*. 2006;23(4):414-420. doi: 10.1093/fampra/cml016.
2. Groeneveld JM, Ballering AV, van Boven K, Akkermans RP, Olde Hartman TC, Uijen AA. Sex differences in incidence of respiratory symptoms and management by general practitioners. *Family practice*. 2020;37(5):631-636. doi: 10.1093/fampra/cmaa040.
3. Arber S, McKinlay J, Adams A, Marceau L, Link C, O'Donnell A. Patient characteristics and inequalities in doctors' diagnostic and management strategies relating to CHD: A video-simulation experiment. *Social Science & Medicine*. 2006;62(1):103-115. doi: 10.1016/j.socscimed.2005.05.028.
4. McKinlay JB, Link CL, Freund KM, Marceau LD, O'Donnell AB, Lutfey KL. Sources of variation in physician adherence with clinical guidelines: Results from a factorial experiment. *Journal of General Internal Medicine*. 2007;22(3):289-296.
5. Kim SE, Paik HY, Yoon H, Lee JE, Kim N, Sung MK. Sex- and gender-specific disparities in colorectal cancer risk. *World Journal of Gastroenterology*. 2015;21(17):5167-5175. doi: 10.3748/wjg.v21.i17.5167.
6. Sourial N, Vedel I, Godard-Sebillotte C, Etches J, Arsenault-Lapierre G, Bronskill SE. Sex differences in dementia primary care performance and health service use: A Population-Based study. *Journal of the American Geriatrics Society*. 2020;68(5):1056-1063. doi: 10.1111/jgs.16347.
7. Ballering AV, Bonvanie IJ, Olde Hartman TC, Monden R, Rosmalen JGM. Gender and sex independently associate with common somatic symptoms and lifetime prevalence of chronic disease. *Social Science & Medicine*. 2020;253:112968. doi: 10.1016/j.socscimed.2020.112968.
8. Ballering AV, Wardenaar KJ, olde Hartman TC, Rosmalen JGM. Female sex and femininity independently associate with common somatic symptom trajectories. *Psychological Medicine*. 2022;52(11):2144-2154. doi: 10.1017/S0033291720004043.
9. Crea F, Battipaglia I, Andreotti F. Sex differences in mechanisms, presentation and management of ischaemic heart disease. *Atherosclerosis*. 2015;241(1):157-168. doi: 10.1016/j.atherosclerosis.2015.04.802.
10. de Jong M, Vos RC, de Ritter R, et al. Sex differences in cardiovascular risk management for people with diabetes in primary care: A cross-sectional study. *BJGP open*. 2019;3(2):bjgpopen19X101645. doi: 10.3399/bjgpopen19X101645.
11. Thompson AE, Anisimowicz Y, Miedema B, Hogg W, Wodchis WP, Aubrey-Bassler K. The influence of gender and other patient characteristics on health care-seeking behaviour: A QUALICOPC study. *BMC family practice*. 2016;17(1):38.
12. Armstrong I, Rochnia N, Harries C, Bundock S, Yorke J. The trajectory to diagnosis with pulmonary arterial hypertension: A qualitative study. *BMJ Open*. 2012;2(2):e000806-000806. Print 2012.
13. Gureje O, Simon GE, Ustun TB, Goldberg DP. Somatization in cross-cultural perspective: A world health organization study in primary care. *The American Journal of Psychiatry*. 1997;154(7):989-995. doi: 10.1176/ajp.154.7.989.
14. Gijsbers van Wijk C, Kolk A, van Den Bosch W, Van Den Hoogen H. Male and female morbidity in general practice: The nature of sex differences. *Social Science & Medicine*. 1992;35(5):665-678. doi: 10.1016/0277-9536(92)90005-b.
15. Halvorsen S, Eritsland J, Abdelnoor M, et al. Gender differences in management and outcome of acute myocardial infarctions treated in 2006-2007. *Cardiology*. 2009;114(2):83-88. doi: 10.1159/000216582.

16. Soler JK, Okkes I. Reasons for encounter and symptom diagnoses: A superior description of patients' problems in contrast to medically unexplained symptoms (MUS). *Family practice*. 2012;29(3):272-282. doi: 10.1093/fampra/cmr101.
17. Luijks H, van Boven K, olde Hartman T, Uijen A, van Weel C, Schers H. Purposeful incorporation of patient narratives in the medical record in the netherlands. *Journal of the American Board of Family Medicine*. 2021;34(4):709-723. doi: 10.3122/jabfm.2021.04.200609.
18. Zijlema WL, Stolk RP, Löwe B, Rief W, White PD, Rosmalen JGM. How to assess common somatic symptoms in large-scale studies: A systematic review of questionnaires. *Journal of Psychosomatic Research*. 2013;74(6):459-68. doi: 10.1016/j.jpsychores.2013.03.093.
19. Acevedo-Mesa A, Tendeiro JN, Roest A, Rosmalen JGM, Monden R. Improving the measurement of functional somatic symptoms with item response theory. *Assessment*. 2020;28(8):1960-1970. doi: 10.1177/1073191120947153.
20. Imai K, Keele L, Tingley D, Yamamoto T. Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *American Political Science Review*. 2011;105(4):765-789.
21. Rosseel Y. Software for mediation analyses. http://www.da.ugent.be/cvs/pages/en/Presentations/Presentation%20Yves%20Rosseel.pdf;. Updated 2013. Accessed October, 2, 2020.
22. Miles J, Shevlin M. Collinearity. In: *Applying regression and correlation: A guide for students and researchers.* Sage; 2001:126-135.
23. Rosseel Y. Lavaan: An R package for structural equation modeling and more. version 0.5-12 (BETA). *Journal of statistical software*. 2012;48(2):1-36.
24. van Boven K, Uijen AA, van de Wiel N, Oskam SK, Schers HJ, Assendelft WJJ. The diagnostic value of the patient's reason for encounter for diagnosing cancer in primary care. *Journal of the American Board of Family Medicine*. 2017;30(6):806-812. doi: 10.3122/jabfm.2017.06.170076 [doi].
25. Vinker S, Kvint I, Erez R, Elhayany A, Kahan E. Effect of the characteristics of family physicians on their utilisation of laboratory tests. *The British Journal of General Practice*. 2007;57(538):377-382.
26. Bertakis KD, Azari R. Patient gender and physician practice style. *Journal of Women's Health*. 2007;16(6):859-868.
27. Hall JA, Roter DL. Do patients talk differently to male and female physicians?: A meta-analytic review. *Patient Education and Counseling*. 2002;48(3):217-224. doi: 10.1016/S0738-3991(02)00174-X.
28. Barsky AJ, Peekna HM, Borus JF. Somatic symptom reporting in women and men. *Journal of General Internal Medicine*. 2001;16(4):266-275. doi: 10.1046/j.1525-1497.2001.00229.x.
29. Hunt K, Adamson J, Hewitt C, Nazareth I. Do women consult more than men? A review of gender and consultation for back pain and headache. *Journal of Health Services Research & Policy*. 2011;16(2):108-117. doi: 10.1258/jhsrp.2010.009131.
30. Clerc Liaudat C, Vaucher P, De Francesco T, et al. Sex/gender bias in the management of chest pain in ambulatory care. *Womens Health*. 2018;14:1745506518805641. doi: 10.1177/1745506518805641.
31. Chakkalakal RJ, Higgins SM, Bernstein LB, et al. Does patient gender impact resident physicians' approach to the cardiac exam? *Journal of General Internal Medicine*. 2013;28(4):561-566. doi: 10.1007/s11606-012-2256-5.
32. Dirkzwager AJ, Verhaak PF. Patients with persistent medically unexplained symptoms in general practice: Characteristics and quality of care. *BMC Family Practice*. 2007;8(1):33. doi: 10.1186/1471-2296-8-33.

33. Evangelidou S, NeMoyer A, Cruz-Gonzalez M, O'Malley I, Alegria M. Racial/ethnic differences in general physical symptoms and medically unexplained physical symptoms: Investigating the role of education. *Cultural Diversity & Ethnic Minority Psychology*. 2020;26(4):557-569. doi: 10.1037/cdp0000319.
34. Brush Jr JE, Krumholz HM, Greene EJ, Dreyer RP. Sex differences in symptom phenotypes among patients with acute myocardial infarction. *Circulation: Cardiovascular Quality and Outcomes*. 2020;13(2):e005948. doi: 10.1161/CIRCOUTCOMES.119.005948.
35. Alabas OA, Tashani OA, Johnson MI. Gender role expectations of pain mediate sex differences in cold pain responses in healthy libyans. *European Journal of Pain*. 2012;16(2):300-311. doi: 10.1016/j.ejpain.2011.05.012.
36. Serpe C. Somatization disorder and gender. In: Nadal KL, ed. *The SAGE encyclopedia of psychology and gender*. SAGE Publications; 2017:1606-1610.
37. Loikas D, Karlsson L, von Euler M, Hallgren K, Schenck-Gustafsson K, Bastholm Rahmner P. Does patient's sex influence treatment in primary care? experiences and expressed knowledge among physicians--a qualitative study. *BMC family practice*. 2015;16:137. doi: 10.1186/s12875-015-0351-5.
38. Nathanson CA. Illness and the feminine role: A theoretical review. *Social Science & Medicine*. 1975;9(2):57-62. doi: 10.1016/0037-7856(75)90094-3.
39. MacLean A, Sweeting H, Hunt K. 'Rules' for boys,'guidelines' for girls: Gender differences in symptom reporting during childhood and adolescence. *Social Science & Medicine*. 2010;70(4):597-604. doi: 10.1016/j.socscimed.2009.10.042.
40. Roter DL, Hall JA, Aoki Y. Physician gender effects in medical communication: A meta-analytic review. *JAMA*. 2002;288(6):756-764. doi: 10.1001/jama.288.6.756.
41. Lutfey KE, Link CL, Marceau LD, et al. Diagnostic certainty as a source of medical practice variation in coronary heart disease: Results from a cross-national experiment of clinical decision making. *Medical Decision Making*. 2009;29(5):606-618. doi: 10.1177/0272989X09331811.
42. Barsky AJ, Orav EJ, Bates DW. Somatization increases medical utilization and costs independent of psychiatric and medical comorbidity. *Archives of General Psychiatry*. 2005;62(8):903-910. doi: 10.1001/archpsyc.62.8.903.
43. Michiels-Corsten M, Donner-Banzhoff N. Beyond accuracy: Hidden motives in diagnostic testing. *Family Practice*. 2018;35(2):222-227. doi: 10.1093/fampra/cmx089.
44. van der Weijden T, van Bokhoven MA, Dinant GJ, van Hasselt CM, Grol RP. Understanding laboratory testing in diagnostic uncertainty: A qualitative study in general practice. *Br J Gen Pract*. 2002;52(485):974-980.
45. Bösner S, Haasenritter Jö, Hani MA, et al. Gender bias revisited: New insights on the differential management of chest pain. *BMC Family Practice*. 2011;12(1):45. doi: 10.1186/1471-2296-12-45.
46. Chapman EN, Kaatz A, Carnes M. Physicians and implicit bias: How doctors may unwittingly perpetuate health care disparities. *Journal of General Internal Medicine*. 2013;28(11):1504-1510. doi: 10.1007/s11606-013-2441-1.
47. Claréus B, Renström EA. Physicians' gender bias in the diagnostic assessment of medically unexplained symptoms and its effect on patient–physician relations. *Scandinavian Journal of Psychology*. 2019;60(4):338-347. doi: 10.1111/sjop.12545.

Appendices

Appendix A: Comorbidities

Malignancies	
ICPC	**Description**
A79	Malignancy NOS
B72	Hodgkin's disease/Lymphoma
B73	Leukaemia
B74	Malig. neoplasm blood other
B75	Beningn/unspecified neoplasm blood
D74	Malig. neoplasm stomach
D75	Malig. neoplasm colon/rectum
D76	Malig. neoplasm pancreas
D77	Malig. Neoplasm digest oth/NOS
F74	Neoplasm of eye/adnexa
H75	Neoplasm of ear
L71	Malig. neoplasm musculoskeletal
N74	Malig. neoplasm nervous system
N75	Benign neoplasm nervous system
N76	Neoplasm nervous system unspec.
R84	Malig. neoplasm bronchus/lung
R85	Malig. neoplasm respiratory, other
R92	Neoplasm respiratory unspecified
T71	Malig. neoplasm thyroid
T73	Neoplasm endocrine oth/unspecified
U75	Malig. neoplasm of kidney
U76	Malig. neoplasm of bladder
U77	Malig. neoplasm urinary other
U79	Neoplasm urinary tract NOS
X75	Malig. neoplasm cervix
X76	Malig. neoplasm breast female
X77	Malig. Neoplasm female genital other
X81	Genital neoplasm female oth/unspec.
Y77	Malig. neoplasm prostate
Y78	Malig. neoplasm male genital other

Cardiovascular disease	
ICPC	**Description**
K72	Cardiovascular neoplasm
K74	Ischaemic heart dis w. angina
K75	Acute myocardial infarction
K76	Ischaemic heart dis w/o angina
K77	Heart failure
K78	Atrial fibrillation/flutter
K79	Paroxysmal tachycardia
K80	Cardiac arrhythmia NOS
K81	Heart/arterial murmur NOS
K82	Pulmonary heart disease
K83	Heart valve disease NOS
K84	Heart disease other
K86	Hypertension uncomplicated
K87	Hypertension complicated
K89	Transient cerebral ischaemia
K90	Stroke/cerebroovascular accident
K91	Cerebrovascular disease
K92	Atherosclerosis/PVD

Asthma and COPD	
ICPC	**Description**
R95	Chronic obstructive pulmonary dis
R96	Asthma

*Abbreviations: "dis" disease, "malig." malignant, "NOS" not otherwise specified, "oth "other, "unspec." Unspecified, "w" with, "w/o" without.

Appendix B. Incidence of the common somatic symptoms as presented at the GP, stratified by sex

	Women (78,954 patient years)	
RFE	**Number of EoC, N (%)**	**Incidence per 1,000 patient years (95%CI)**
Headache	2,153 (10.2)	27.3 (26.1-28.4)
Dizziness	1,782 (8.5)	22.6 (21.5-23.6)
Heart pain	337 (1.6)	4.3 (3.8-4.7)
(Lower) back pain	3,485 (16.6)	44.1 (42.7-45.6)
Nausea	901 (4.3)	11.4 (10.7-12.2)
Muscle pain	322 (1.5)	4.1 (3.6-4.5)
Shortness of breath	2,103 (10.0)	26.6 (25.5-27.8)
Chills	67 (0.3)	0.9 (0.7-1.1)
Tingling fingers, feet and/or toes	365 (1.7)	4.6 (4.2-5.1)
Swallowing/throat symptoms	3,034 (14.4)	38.4 (37.1-39.8)
Weakness/general tiredness	3,694 (17.6)	46.8 (45.3-48.3)
Arm or leg symptoms	2,788 (13.3)	35.3 (34.0-36.6)
Total	21,031 (100.0)	266.4 (263.3-269.5)

Appendix C. Multilevel multiple logistic regression analyses: associations stratified per RFE between the patients' sex and the interventions conducted during an EoC that started with a common somatic symptom.

Interventions	**Number of EoC (%)**	
	Women (%)	Men (%)
Interventions for chills (A02)		
Physical examination	39 (58.2)	54 (83.1)
Laboratory diagnostics	20 (29.9)	16 (24.6)
Diagnostic imaging	1 (1.4%)	1 (1.5%)
Referral to specialist	10 (14.9)	14 (21.5)
Interventions for weakness/general tiredness (A04)		
Physical examination	2,599 (70.4)	1,467 (76.6)
Laboratory diagnostics	2,759 (74.7)	1,311 (68.5)
Diagnostic imaging	104 (2.8)	73 (3.8)
Referral to specialist	284 (7.7)	200 (10.4)
Interventions for nausea (D09)		
Physical examination	568 (63.0)	223 (69.5)
Laboratory diagnostics	203 (22.5)	69 (21.5)
Diagnostic imaging	34 (3.8)	10 (3.1)
Referral to specialist	86 (9.5)	45 (14.0)

Men (75,734 patient years)	
Number of EoC, N (%)	**Incidence per 1,000 patient years (95%CI)**
1,258 (9.5)	16.6 (15.7-17.5)
927 (7.0)	12.2 (11.5-13.0)
323 (2.4)	4.3 (3.8-4.7)
2,477 (18.7)	32.7 (31.4-34.0)
321 (2.4)	4.2 (3.8-4.7)
232 (1.8)	3.1 (2.7-3.6)
1,490 (11.3)	19.7 (18.7-20.7)
65 (0.5)	0.9 (0.7-1.1)
244 (1.8)	3.2 (2.8-3.6)
1,927 (14.6)	25.4 (24.3-26.6)
1,914 (14.5)	25.3 (24.2-26.4)
2,059 (15.6)	27.2 (26.0-28.4)
13,237 (100.0)	174.8 (172.1 – 177.5)

Number of EoC (%)	**Odds ratio**	
Total	Unadjusted	Adjusted[a]
93 (70.5)	0.28 (0.13-0.64)	**0.26 (0.10-0.66)**
36 (27.3)	1.30 (0.60-2.81)	1.54 (0.58-4.06)
2 (1.5%)	0.97 (0.60-15.8)	0.91 (0.15-5.55)
24 (18.2)	0.64 (0.26-1.56)	0.68 (0.25-1.85)
4,066 (72.5)	0.72 (0.64-0.82)	**0.72 (0.63-0.82)**
4,070 (72.6)	1.36 (1.20-1.53)	**1.37 (1.21-1.55)**
177 (3.2)	0.73 (0.54-0.99)	0.86 (0.66-1.11)
484 (8.6)	0.71 (0.59-0.86)	**0.74 (0.61-0.90)**
791 (64.7)	0.75 (0.57-0.99)	0.79 (0.57-1.07)
272 (22.2)	1.06 (0.78-1.45)	1.14 (0.81-1.59)
44 (3.6)	1.22 (0.60-2.50)	1.11 (0.61-2.02)
131 (10.7)	0.65 (0.44-0.95)	0.67 (0.44-1.01)

Appendix C. Continued.

Interventions	Number of EoC (%)	
	Women (%)	Men (%)
Interventions for swallowing/throat symptoms (D21/R21)		
Physical examination	2,498 (82.3)	1,631 (84.6)
Laboratory diagnostics	383 (12.6)	211 (10.9)
Diagnostic imaging	36 (1.2)	25 (1.3)
Referral to specialist	243 (8.0)	165 (8.6)
Interventions for heart pain (K01)		
Physical examination	261 (77.4)	261 (80.8)
Laboratory diagnostics	73 (21.7)	66 (20.4)
Diagnostic imaging	6 (1.8)	7 (2.2)
Referral to specialist	114 (3.8)	134 (41.5)
Interventions for (lower) back pain (L02/L03)		
Physical examination	2,620 (75.2)	1,938 (78.2)
Laboratory diagnostics	359 (10.3)	153 (6.2)
Diagnostic imaging	502 (14.4)	339 (13.7)
Referral to specialist	244 (7.0)	180 (7.3)
Interventions for arm and leg symptoms (L09/L14)		
Physical examination	2,408 (86.4)	1,797 (87.2)
Laboratory diagnostics	310 (11.1)	184 (8.9)
Diagnostic imaging	368 (13.2)	296 (14.4)
Referral to specialist	387 (13.9)	299 (14.5)
Interventions for muscle pain (L18)		
Physical examination	219 (68.0)	171 (73.7)
Laboratory diagnostics	98 (30.4)	79 (34.1)
Diagnostic imaging	11 (3.4)	13 (5.6)
Referral to specialist	36 (11.2)	22 (9.5)
Interventions for headache (N01)		
Physical examination	1,615 (75.0)	977 (77.6)
Laboratory diagnostics	255 (11.8)	154 (12.2)
Diagnostic imaging	30 (1.4)	21 (1.7)
Referral to specialist	223 (10.4)	141 (11.2)
Interventions for tingling in fingers, feet and/or toes (N05)		
Physical examination	311 (85.2)	204 (83.6)
Laboratory diagnostics	72 (19.7)	55 (22.5)
Diagnostic imaging	3 (0.8)	6 (2.5)
Referral to specialist	61 (16.7)	33 (13.5)

Number of EoC (%)	Odds ratio	
Total	Unadjusted	Adjusted[a]
4,129 (83.2)	0.85 (0.72-0.99)	0.97 (0.80-1.18)
594 (12.0)	1.18 (0.98-1.41)	**1.21 (1.00-1.45)**
61 (1.2)	0.91 (0.55-1.53)	0.99 (0.73-1.35)
408 (8.2)	0.93 (0.76-1.14)	0.98 (0.79-1.20)
522 (79.1)	0.82 (0.56-1.19)	1.03 (0.67-1.60)
660 (21.1)	1.08 (0.74-1.57)	1.27 (0.84-1.91)
13 (2.0)	0.82 (0.27-2.46)	0.99 (0.45-2.17)
248 (37.6)	0.72 (0.53-0.99)	0.77 (0.54-1.10)
4,558 (76.5)	0.84 (0.75-0.95)	**0.78 (0.68-0.90)**
512 (8.6)	1.75 (1.43-2.12)	**1.56 (1.29-1.90)**
841 (14.1)	1.06 (0.92-1.23)	1.02 (0.88-1.20)
424 (7.1)	0.96 (0.79-1.17)	0.92 (0.75-1.13)
4,205 (86.8)	0.93 (0.79-1.10)	1.01 (0.84-1.21)
494 (10.2)	1.28 (1.05-1.55)	1.20 (0.99-1.46)
552 (11.4)	0.91 (0.77-1.07)	0.95 (0.80-1.12)
686 (14.2)	0.95 (0.81-1.12)	0.93 (0.78-1.11)
390 (70.4)	0.76 (0.52-1.10)	0.78 (0.52-1.18)
177 (31.9)	0.85 (0.59-1.22)	0.86 (0.58-1.28)
24 (4.3)	0.60 (0.26-1.36)	0.76 (0.37-1.57)
58 (10.5)	1.20 (0.69-2.10)	1.23 (0.68-2.26)
2,592 (76.0)	0.87 (0.74-1.02)	0.84 (0.70-1.01)
409 (12.0)	0.97 (0.78-1.19)	0.92 (0.74-1.15)
51 (1.5)	0.83 (0.48-1.46)	0.94 (0.65-1.35)
364 (10.7)	0.92 (0.73-1.15)	0.84 (0.66-1.06)
515 (84.6)	1.13 (0.73-1.76)	1.36 (0.84-2.21)
127 (20.9)	0.84 (0.57-1.25)	0.84 (0.56-1.28)
9 (1.5)	0.33 (0.08-1.33)	0.75 (0.32-1.77)
94 (15.4)	1.28 (0.81-2.03)	1.29 (0.79-2.08)

Appendix C. Continued.

Interventions	Number of EoC (%)	
	Women (%)	Men (%)
Interventions for dizziness (N17)		
Physical examination	1,461 (82.0)	778 (83.8)
Laboratory diagnostics	558 (31.3)	201 (21.7)
Diagnostic imaging	10 (0.6)	9 (1.0)
Referral to specialist	170 (9.5)	115 (12.4)
Interventions for shortness of breath (R02)		
Physical examination	1,903 (90.6)	1,375 (92.1)
Laboratory diagnostics	559 (26.6)	376 (25.2)
Diagnostic imaging	159 (7.6)	151 (10.1)
Referral to specialist	302 (14.4)	263 (17.6)

[a] Adjusted for: age, number of contacts in an EoC, type of consult and comorbidities at the start of an EoC. Odds ratios reflect women compared to men.

Appendix D. Top 5 diagnoses for common somatic symptoms divided by sex.

Top 5 final diagnoses per RFE (ICPC)[a]	Women, N (%)	Men, N (%)
Episodes of care starting with chills (ICPC=A02; N=132)		
Chills (A02)	18 (26.9)	15 (23.1)
Urinary tract Infection (U71)	9 (13.4)	8 (12.3)
Pneumonia (R81)	5 (7.5)	8 (12.3)
Fever (A03)	4 (6.0)	7 (10.8)
Upper respiratory infection acute (R74)	3 (4.5)	5 (7.7)
Episodes of care starting with weakness/general tiredness (ICPC=A04; N=5,608)		
Weakness (A04)	2,612 (70.7)	1,312 (68.5)
Upper respiratory infection acute (R74)	132 (3.6)	86 (4.5)
Surmenage/Neuraesthenia (P78)	57 (1.5)	29 (1.5)
Pneumonia (R81)	46 (1.2)	37 (1.9)
Influenza (R80)	29 (0.8)	33 (1.7)
Episodes of care starting with nausea (ICPC=D09; N=1,222)		
Nausea (D09)	355 (39.4)	120 (37.4)
Gastroenteritis presumed infection (D73)	67 (7.4)	28 (8.7)
Adverse effect medical agent (A85)	39 (4.3)	17 (5.3)
Dizziness (N17)	23 (2.6)	8 (2.5)
Abdominal pain localized other (D06)	19 (2.1)	10 (3.1)

Number of EoC (%)	Odds ratio	
Total	Unadjusted	Adjusted[a]
2,239 (82.7)	0.88 (0.71-1.09)	0.93 (0.74-1.17)
759 (28.0)	1.65 (1.37-1.99)	**1.68 (1.38-2.04)**
19 (0.7)	0.58 (0.23-1.42)	0.94 (0.60-1.47)
285 (10.5)	0.75 (0.58-0.96)	0.80 (0.61-1.04)
3,278 (91.2)	0.83 (0.65-1.05)	0.87 (0.68-1.10)
935 (26.0)	1.08 (0.93-1.25)	1.04 (0.88-1.23)
310 (8.6)	0.73 (0.58-0.92)	**0.72 (0.58-0.92)**
565 (15.7)	0.79 (0.66-0.94)	**0.78 (0.65-0.95)**

Appendix D. Continued

Top 5 final diagnoses per RFE (ICPC)[a]	Women, N (%)	Men, N (%)
Episodes of care starting with swallowing/throat symptoms (ICPC=D21/R21; N=4,961)		
Throat symptoms (R21)	981 (32.3)	625 (32.4)
Upper respiratory infection acute (R74)	822 (27.1)	509 (26.4)
Tonsillitis acute (R76)	601 (19.8)	335 (17.4)
Swallowing problem (D21)	97 (3.2)	76 (3.9)
Hypertrophy tonsils/adenoids (R90)	52 (1.7)	29 (1.5)
Episodes of care starting with heart pain (ICPC=K01; N=660)		
Heart pain (K01)	153 (45.4)	126 (39.0)
Chest symptom/complaint (L04)	60 (17.8)	59 (18.3)
Acute myocardial infarction (K75)	6 (1.8)	24 (7.4)
Ischaemic heart disease with Angina (K74)	5 (1.5)	20 (6.2)
Chest pain NOS (A11)	13 (3.9)	7 (2.2)
Episodes of care starting with (lower) back pain (ICPC=L02/L03; N=5,962)		
Low back symptom/complaint (L03)	1,658 (47.6)	1,409 (56.9)
Back symptom/complaint (L02)	819 (23.5)	546 (22.0)
Back syndrome with radiating pain (L86)	167 (4.8)	153 (6.2)
Urinary tract infection (U71)	133 (3.8)	13 (0.5)
Back syndrome w/o radiating pain (L84)	64 (1.8)	46 (1.9)

Appendix D. Continued

Top 5 final diagnoses per RFE (ICPC)[a]	Women, N (%)	Men, N (%)
Episodes of care starting with arm or leg symptoms (ICPC=L09/L14; N=4,847)		
Leg symptom/complaint (L14)	896 (32.1)	647 (31.4)
Arm symptom/complaint (L09)	327 (11.7)	243 (11.8)
Bursitis/tendinitis/synovitis NOS (L87)	189 (6.8)	104 (5.0)
Injury musculoskeletal NOS (L81)	109 (3.9)	118 (5.7)
Bruise/contusion (S16)	99 (3.6)	101 (4.9)
Episodes of care starting with muscle pain (ICPC=L18; N=554)		
Muscle pain (L18)	245 (76.1)	157 (67.7)
Adverse effect medical agent (A85)	13 (4.0)	14 (6.0)
Influenza (R80)	5 (1.6)	12 (5.2)
Upper respiratory tract infection (R74)	6 (1.9)	6 (2.6)
Muscle symptom/complaint NOS (L19)	6 (1.9)	4 (1.7)
Episodes of care starting with headache (ICPC=N01; N=3,411)		
Headache (N01)	1,006 (46.7)	589 (46.8)
Sinusitis acute/chronic (R75)	198 (9.2)	111 (8.8)
Tension headache (N95)	195 (9.1)	87 (6.9)
Migraine (N89)	115 (5.3)	44 (3.5)
Upper respiratory tract infection (R74)	85 (3.9)	50 (4.0)
Episodes of care starting with tingling of fingers, feet and/or toes (ICPC=N05; N=609)		
Tingling of Fingers/feet/toes (N05)	179 (49.0)	108 (44.3)
Carpal tunnel syndrome (N93)	71 (19.5)	32 (13.1)
Peripheral neuritis/neuropathy (N94)	25 (6.8)	35 (14.3)
Neck symptom/complaint (L01)	12 (3.3)	7 (2.9)
Hyperventilation syndrome (R98)	9 (2.5)	8 (3.3)
Episodes of care starting with vertigo/dizziness (ICPC=N17; N=2,709)		
Dizziness (N17)	991 (55.6)	459 (49.5)
Vertiginous syndrome (H82)	367 (20.6)	197 (21.2)
Postural hypotension (K88)	59 (3.3)	40 (4.3)
Headache (N01)	16 (0.9)	19 (2.0)
Fainting/Syncope (A06)	18 (1.0)	14 (1.5)
Episodes of care starting with shortness of breath/dyspnoea (ICPC=R02; N=3,593)		
Shortness of breath (R02)	606 (28.9)	361 (4.2)
Upper respiratory tract infection (R74)	401 (19.1)	266 (17.8)
Pneumonia (R81)	180 (8.6)	156 (10.4)
Acute bronchitis/bronchiolitis (R78)	148 (7.0)	128 (8.6)
Asthma (R96)	89 (4.2)	66 (4.4)

[a]Final diagnoses include both symptom diagnoses (operationalized by an ICPC-code ≤30, in which symptom remain symptoms over the course of an EoC) and disease diagnoses (operationalized by an ICPC-code ≥70, in which symptoms evolve in a diagnosed disease over the course of an EoC).

Appendix E. Mediation effects of diagnostic interventions on the association between female sex and disease diagnosis, expressed by probit coefficients.

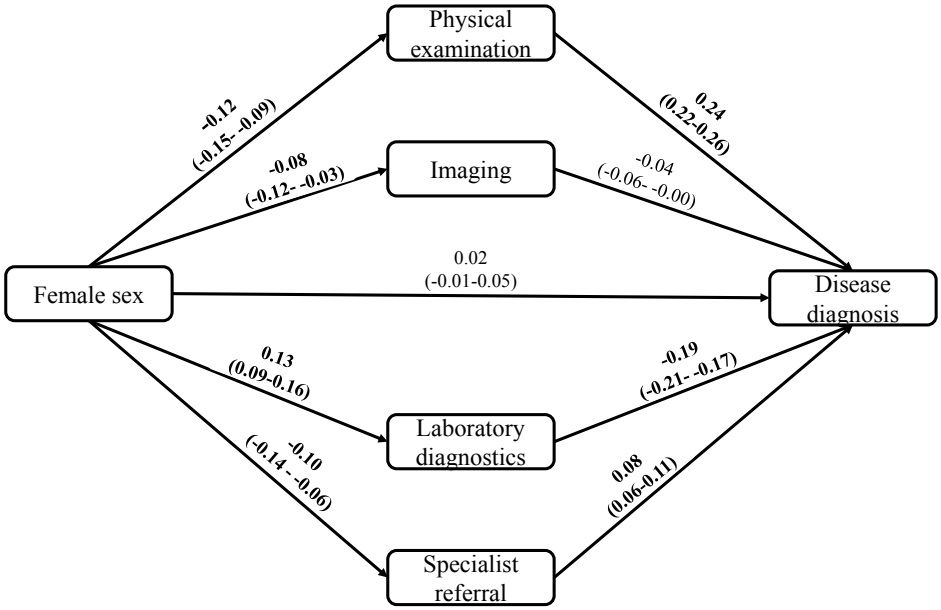


Figure E: Mediation effects

CHAPTER 10

Ballering, A.V., Rosmalen, J.G.M., & olde Hartman, T.C. (2022). Differences between women and men are present in the rate of diagnosed diseases after a diagnostic intervention is conducted in primary care. *The Journal of the American Board of Family Medicine*, *35*(1), 73-84.

Abstract

Background: Recently it was shown that the relative lack of diagnostic interventions conducted in women mediated the negative association between female sex and diagnosed disease. However, it remains unknown whether women and men receive disease diagnoses in an equal frequency after diagnostic interventions have been performed in general practice.

Methods: We used generalized linear mixed-effect models to assess the association between diagnostic interventions and disease diagnoses when patients presented with common somatic symptoms and studied whether the association differed between female and male patients.

Results: In 34,268 episodes of care (61.4% female) physical examinations and specialist referrals were associated with more disease diagnoses (OR = 2.32; 95% CI = 2.17–2.49 and OR = 1.38; 95% CI = 1.27–1.49, respectively), whereas laboratory diagnostics were associated with fewer disease diagnoses (OR = 0.50; 95% CI = 0.47–0.54). Significant interaction terms showed that women presenting with back pain, tiredness, arm and/or leg symptoms and tingling extremities were provided with fewer disease diagnoses after diagnostic interventions were performed than men. We found no significant interaction term that indicated that men were provided with fewer disease diagnoses after a diagnostic intervention than women.

Conclusion: Especially when patients present with the mentioned symptoms, general practitioners should be aware that diagnostic interventions yield fewer disease diagnoses in female patients than in men. Yet, performing fewer diagnostic interventions in women with these symptoms will further exacerbate sex differences in disease diagnoses.

Background

Multiple studies have shown that female patients' common somatic symptoms remain more often medically unexplained than male patients' somatic symptoms.[1,2] Yet, only recently it has been shown that the negative association between female sex and medically unexplained symptoms is mediated by the performed diagnostic interventions in primary care.[3]

Recent studies show that differences between women and men in primary care diagnostic trajectories for somatic symptoms are not uncommon. A study that followed patients from common somatic complaint presentation to final diagnosis has shown that women receive fewer physical examinations, less diagnostic imaging and fewer referrals to a specialist than men when they present with common somatic symptoms in general practice.[3] A similar pattern is observed when patients present with cough and/or shortness of breath in general practice.[4] In addition, female patients who present symptoms suggestive of coronary heart disease, have a lower likelihood of receiving a physical examination that follows the guidelines and of being referred to a cardiologist than male patients.[5-7]

Little to no research focused on whether these sex differences in the rate of diagnostic interventions in patients presenting somatic symptoms are justified, as it has not yet been studied whether women and men receive disease diagnoses in an equal frequency after diagnostic interventions have been performed in general practice. In other words, it remains unknown whether a diagnostic intervention in general practice, such as physical examinations, diagnostic imaging, laboratory diagnostics or referrals to a specialist, associates differently in women and men with a disease diagnosis (ie, with explained symptoms).

Clinically, to ensure equal and appropriate care for all patients, irrespective of a patient's sex, it is pivotal for general practitioners (GPs) to be aware of whether the odds of receiving a disease diagnosis after a certain diagnostic intervention differs significantly between male and female patients presenting with somatic symptoms. Therefore, the aim of this study is to assess whether differences between men and women are present in the association between diagnostic interventions and a disease diagnosis when patients present themselves with common somatic symptoms. In this study data derived from the Dutch Family Medicine Network (FaMe-Net), a practice-based research network, is analyzed by generalized linear mixed-effect models.

Methods

Study Design

Our study included data from FaMe-Net, in which approximately 32,000 patients from 26 GPs working in seven general practices throughout the Netherlands are included[8]. FaMe-Net is the world's oldest practice-based research network and has registered patients' morbidity in an episode of care (EoC) structure since its inception. GPs systematically code all information, including reason for encounter (RFE), interventions (ie, physical examinations, laboratory diagnostics, diagnostic imaging and referrals to specialists), and final diagnoses, according to the International Classification of Primary Care (ICPC-2). An EoC is defined as a patient's health problem from the first encounter with the GP until the last encounter related to that specific health problem.

Although an EoC can start with multiple RFEs, one final diagnosis is ultimately linked to all encounters within an EoC. A final diagnosis could be either a disease diagnosis or symptom diagnosis. A disease diagnosis was defined as symptoms, followed over time, that evolve in a diagnosed disease (operationalized as ICPC≥70, including psychiatric ICPC codes), whereas a symptom diagnosis was defined as symptoms, when followed over time, that continued to be symptoms as relevant diagnostic criteria were not met (operationalized as ICPC≤30). The RFE of an EoC should be acknowledged by patients as an adequate description of their demand of care and can be a symptom, a self-diagnosis, or a request for a particular intervention. The quality of data registration within FaMe-Net is high, as participating GPs regularly meet to discuss registration logistics and diagnostic criteria[8]. Moreover, the automated GP information system recognizes errors and inconsistencies in registration.

For this study, we selected EoCs that started with a common somatic symptom between January 1, 2014 and December 31, 2018. Contacts within an EoC that continued after this date were not included. Face-to-face encounters as well as telephone, and digital consultations were included. We included 15 RFEs related to 12 symptoms: headache (ICPC-N01), dizziness (ICPC-N17), heart pain (ICPC-K01), (lower) back pain (ICPC-L02 and ICPC-L03), nausea (ICPC-D09), muscle pain (ICPC-L18), shortness of breath/dyspnea (ICPC-R02), chills (ICPC-A02), tingling of fingers, feet, and/or toes (ICPC-N05), swallowing/throat problems (ICPC-D21 and ICPC-R21), weakness or general tiredness (ICPC-A04), and arm or leg symptoms (ICPC-L09 and ICPC-L14). These 12 symptoms reflect the contents of the Symptom Checklist-90 Somatization subscale (SCL-90 SOM),[9] are common,[10] and often remain unexplained.[11] All EoC that started with the same RFE on the same date within the same patient were excluded (n=106). When the EoC started with more than one of the included RFEs (n=1,605; 4.7%) we analyzed the first-mentioned RFE.

Statistical Analyses

To assess whether diagnostic interventions, namely physical examinations (ICPC-30 and ICPC-31), laboratory diagnostic interventions (ICPC-33 and ICPC-34), diagnostic imaging (ICPC-41) and specialist referrals (ICPC-67) were associated with a disease diagnosis, we defined generalized linear mixed-effect models. EoCs are nested within an individual, thus we clustered analyses at the patient level. Patients' sex, patients' age at time of diagnosis, the number of contacts between a patient and GP during an EoC, type of consult (face-to-face, digital, or by phone), the type of RFE, the presence of comorbidities at the start of an EoC (***Appendix A***) and the four aforementioned diagnostic interventions were included as independent variables. In addition, we included interaction terms between patient's sex and the diagnostic interventions to assess whether the association between the respective diagnostic intervention and disease diagnosis differed between female and male patients.

To test whether the continuous covariates (patients' age and number of contacts between a patient and GP within an EoC) included in the analyses fulfilled the linearity assumption of multiple logistic regression, we divided the covariates into categories, and assessed whether the estimates changed monotonically. In addition, we found no indication for multicollinearity, as the variance inflation factor was <5 in all analyses.[12] The statistical analyses and descriptive analyses, including chi square tests, Mann Whitney U tests, and independent T-tests were conducted in IBM SPSS Statistics v. 25. We maintained a two-sided a-value of $p<.01$ to correct for multiple testing. We adhered to the STROBE and SAGER guidelines for reporting observational cohort studies.[13]

Results

We identified 34,268 EoCs that started with a common somatic symptom in 10,541 female patients and 7,915 male patients. The majority of these EoC started with 1 RFE (71.7%) and most EoCs only involved 1 encounter with the GP (65.1%). A more detailed overview of the study population is provided in ***Table 1***.

Sex Differences in Performed Diagnostic Interventions

Figure 1 shows that three diagnostic interventions performed by the GP to obtain a diagnosis are significantly more often performed per EoC in male than in female patients. Only laboratory diagnostics are more often performed in female patients. ***Appendix B*** shows the frequencies of the performed interventions stratified by sex and RFE.

Table 1: Overview of the study population

Characteristic		Female	Male	*p*-value
Patients, N		10,541 (57.1%)	7,915 (42.9%)	
Patients' age in years, Mean (SD)		43.4 (23.1)	42.3 (23.9)	**<0.001**§
Patient years*, N		78,954 (52.3%)	75,734 (47.7%)	
Unique EoC, N		21,025 (61.4%)	13,243 (38.6%)	
EoC with intervention	No intervention	3,218 (15.3%)	1,827 (13.8)	
	≥1 intervention	17,807 (84.7%)	11,416 (86.2%)	**<0.001**±
EoC with final diagnosis	Disease diagnosis	7,470 (35.5%)	4,924 (37.2%)	**0.002**±
	Symptom diagnosis	13,555 (64.5%)	8,319 (62.8%)	
Encounters within an EoC	1 encounter	13,619 (64.8%)	8,697 (65.7%)	0.090±
	>1 encounter	7,406 (35.2%)	4,546 (34.3%)	
RFE per EoC	1 RFE	14,739 (70.1%)	9,839 (74.3%)	**<0.001**±
	>1 RFE	6,286 (29.9%)	3,404 (25.7%)	
EoC with comorbidities, N		8,282 (39.4%)	5,295 (43.7%)	0.271±
Cardiovascular disease		5,741 (27.3%)	3,827 (28.9%)	**0.001**±
Asthma and/or COPD		3,106 (14.8%)	1,998 (15.1%)	0.427±
Malignancies		1,684 (8.0%)	979 (7.4%)	0.038±

*The cumulative years included patients were at risk of an incident common somatic symptom; ±Chi-square test; ‡Mann-Whitney U test; §Independent T-test

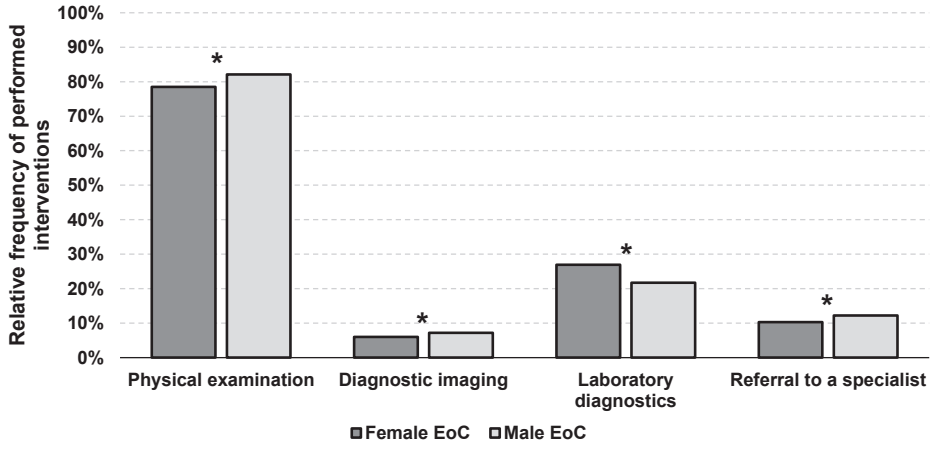


Figure 1. Relative frequency of the performed interventions in female and male episodes of care related to common somatic symptoms. (*p<0.001, chi square test).

Sex differences in Outcomes of Diagnostic Interventions

Table 2 shows that the odds of receiving a disease diagnosis were signficantly higher in EoCs in which a physical examination was performed (odds ratio [OR]=2.32, 95%CI=2.17-2.49) or when the GP requested a referral to a specialist (OR=1.38, 95%CI=1.27-1.49). In contrast, a request for laboratory diagnostics increased the odds of a symptom diagnosis (OR=0.50, 95%CI=0.47-0.54).

We found a significant interaction term between sex and laboratory diagnostics (OR=0.82, 95%CI=0.73-0.93), as well as a significant sex-by-specialist referral interaction term (OR=0.84, 95%CI=0.72-0.97). These estimates indicate that if a GP requests laboratory diagnostics or a specialist referral, the odds of receiving a disease diagnosis are significantly lower for female patients than for male patients.

Table 2: Adjusted associations between diagnostic interventions and disease diagnoses in episodes of care starting with common somatic symptoms, stratified by sex.

		OR (95%CI)*	
Independent variables	**Total EoC (N=34,268)**	**Female EoC (N=21,025)**	**Male EoC (N=13,243)**
Physical examination	**2.32 (2.17-2.49)**	**2.33 (2.14-2.54)**	**2.32 (2.07-2.60)**
Laboratory diagnostics±	**0.50 (0.47-0.54)**	**0.47 (0.44-0.52)**	**0.56 (0.50-0.62)**
Imaging	0.96 (0.86-1.06)	0.92 (0.81-1.06)	1.00 (0.85-1.17)
Referral to a specialist±	**1.38 (1.27-1.49)**	**1.29 (1.17-1.43)**	**1.50 (1.33-1.69)**

*Adjusted for patients' sex (total EoC), patients' age, presence of comorbidities at the start of an EoC, number of contacts within an EoC, type of consult, and type of RFE. ±The interaction term between the patient's sex and the respective diagnostic intervention was statistically significant.

Sex Differences in Outcomes of Diagnostic Interventions per RFE

Table 3 shows in more detail which diagnostic interventions associate with a disease diagnosis, stratified per RFE. For example, in (lower) back pain, it was found that all diagnostic interventions, except for imaging, are associated with a disease diagnosis. In addition, in the patients with (lower) back pain we found significant interaction terms between patients' sex and physical examinations (OR=0.51, 95%CI=0.34-0.78), patients' sex and imaging (OR=0.64, 95%CI=0.43-0.96), and patients' sex and a specialist referral (OR=0.46, 95%CI=0.29-0.72). This indicates that female patients with lower back pain received significantly fewer disease diagnoses after they underwent these diagnostic interventions than their male counterparts.

Table 3: Adjusted odds ratios and significant intervention-by-female sex interaction terms for the association between diagnostic interventions and disease diagnosis, stratified by RFE.

	OR* (95% CI) **Headache (N=3,411)**
Physical examination	**2.07 (1.74-2.47)**
Laboratory diagnostics	**0.59 (0.47-0.74)**
Imaging	0.67 (0.37-1.22)
Specialist referral	**0.77 (0.60-0.98)**
	OR* (95% CI) **Nausea (N=1,222)**
Physical examination	**1.93 (1.46-2.56)**
Laboratory diagnostics	**0.49 (0.35-0.68)**
Imaging	0.57 (0.26-1.21)
Specialist referral	1.25 (0.82-1.90)
	OR* (95% CI) **Tingling in fingers, feet and/or toes (N=609)**
Physical examination	1.16 (0.72-1.87)
Laboratory diagnostics	**0.40 (0.25-0.64)**±
Imaging	0.47 (0.11-2.06)
Specialist referral	**1.72 (1.05-2.81)**

*Adjusted for the patients' age at diagnosis, comorbidities, number of contacts per EoC and sex.
±Significant intervention-by-female sex interaction term.

Similarly, in EoCs starting with tingling sensations in extremities, with tiredness, or with arm and/or leg symptoms, we found that receiving laboratory diagnostics was associated with fewer disease diagnoses (**Table 3**). In these EoCs, significant interaction terms between patients' sex and laboratory diagnostics (OR=0.32, 95%CI=0.13-0.81; OR=0.72, 95%CI=0.53-0.96; OR=0.56 95%CI=0.37-0.86, respectively) were found as well ***(Appendix C)***. These indicate that female patients who were provided with laboratory diagnostics upon presenting these complaints, less frequently received a disease diagnosis than male patients with these complaints that received laboratory diagnostics. All significant interaction terms between female sex and the respective diagnostic interventions indicated that women are less likely diagnosed with a disease after being provided with a diagnostic intervention than men. An overview of all statistically significant interaction terms between sex and a diagnostic intervention stratified by RFE is given in ***Appendix C***.

	OR* (95% CI)	
Dizziness (N=2,709)	**Heart pain (N=660)**	**(Lower) back pain (N=5,962)**
2.12 (1.68-2.67)	0.98 (0.59-1.64)	**1.35 (1.11-1.63)**±
0.30 (0.25-0.37)	0.94 (0.56-1.58)	**1.49 (1.18-1.88)**
0.43 (0.14-1.36)	0.51 (0.12-2.23)	0.87 (0.70-1.07)±
1.11 (0.84-1.46)	1.14 (0.73-1.77)±	**2.78 (2.18-3.55)**±
	OR* (95% CI)	
Muscle pain (N=554)	**Shortness of breath (N=3,593)**	**Chills (N=132)**
1.19 (0.71-2.01)	**2.58 (2.02-3.31)**	**3.35 (1.23-9.11)**
0.71 (0.41-1.23)	**0.78 (0.65-0.92)**	0.81 (0.29-2.28)
0.30 (0.06-1.39)	**0.49 (0.38-0.64)**	0.69 (0.02-19.9)
0.66 (0.27-1.59)	1.07 (0.87-1.31)	0.57 (0.19-1.74)
	OR* (95% CI)	
Swallowing/throat symptoms (N=4,961)	**Weakness/general tiredness (N=5,608)**	**Arm and/or leg symptoms (N=4,847)**
5.42 (4.56-6.44)	**2.57 (2.14-3.09)**	**1.70 (1.40-2.06)**
0.97 (0.80-1.18)	**0.20 (0.18-0.24)**±	**0.51 (0.41-0.63)**±
0.40 (0.23-0.72)	**1.62 (1.14-2.32)**	1.11 (0.93-1.32)
0.64 (0.51-0.81)	1.25 (0.99-1.58)	**1.86 (1.55-2.23)**±

Discussion

This study is the first to show that diagnostic interventions are differently associated with disease diagnoses in female and male patients presenting with common somatic symptoms in general practice. Women receive fewer disease diagnoses than men after a diagnostic intervention, for example in (lower) back pain men more often receive a disease diagnosis after a physical examination, imaging and a specialist referral than women. Similarly, when laboratory diagnostics are performed, male patients with tingling fingers, feet and/or toes, tiredness, and arm and/or leg symptoms more often receive a disease diagnosis than female patients.

Limitations and Strengths

Results of this study should be interpreted in the light of its limitations. First, patients' full medical history was unknown, and the adjustment for comorbidities is non-exhaustive as it included asthma/chronic obstructive pulmonary disease, cardiovascular disease, and malignancies. Second, not all patient characteristics that may affect the association between diagnostic interventions and final diagnosis were known, including patients' socioeconomic status, ethnicity and time before seeking help. Third, the sex and gender of the involved GPs is unknown. Concordance between a patient's and GP's sex is important in patients' diagnostic trajectories. Female physicians, for instance, are more inclined to conduct female preventive procedures than male physicians.[14,15] Male patients have also been found to be more demanding and assertive toward their GP than female patients,[16] an effect that is amplified in case of female GPs.[17]

This study also has strengths. First, the data were extracted from a long-lasting, large primary care registration network, which minimizes the risk of recall bias. FaMe-Net is the only existing database in which GPs register the RFE, diagnoses and conducted interventions for all contacts within an EoC. The data are detailed, valid, and accurate, as participating GPs frequently discuss diagnostic coding[8]. Furthermore, the use of RFEs to identify EoCs related to common somatic symptom reduces bias in the data, as it avoids the GPs' interpretation of a patient's complaint.[18]

Comparison to Literature

In line with previous research, we found that fewer diagnostic interventions were conducted in female EoCs, except for laboratory diagnostics.[4-6] For example, one of these former studies focusing on cough and dyspnea, showed that women with these complaints are less likely to be provided with diagnostic interventions and a disease diagnosis compared with their male counterparts. This could be due to possible sex-related differences in help-seeking behavior.[19] Seeking help early in the disease process may result in women presenting less typical symptoms for which the GP does not perform or request diagnostic interventions. Ultimately, this may result in underdiagnosis in female patients. However, the diagnosis of a disease is not necessarily a direct indicator of improved care.

In addition, it has been argued that fewer physical examinations in female patients may be due to GPs' concerns about the shame that female patients may experience during a physical examination, leading to hesitance in GPs to examine, for example, intimate areas.[20] Others have argued that the abundance of male-bodied images in medical anatomy textbooks results in a sex bias in physical examinations.[21]

However, in female EoCs more laboratory diagnostics were ordered than in male EoCs. This may be due to uncertainty, either in the GP or in the patient. The GP may experience diagnostic uncertainty in female patients, which may interact with anticipated regret of missing a serious disease.[22,23] Especially laboratory diagnostics were found to associate with mitigating diagnostic uncertainty in GPs.[23] Notably, previous research on sex differences in diagnostic uncertainty of the GP focused on cardiovascular disease and found a greater diagnostic uncertainty among GPs when assessing female patients[22,24]; whether this occurs in common somatic symptoms in general practice as well has not been investigated yet. Female patients may experience uncertainty about the explanation of symptoms given by the GP, possibly due to their heightened tendency to ruminate.[25] As laboratory diagnostic tests are often used as a means to reassure patients, and more easily applied than imaging or specialist referrals,[23] GPs may be more inclined to order laboratory diagnostic tests to reassure female patients.

We also show that women received fewer disease diagnoses than men after diagnostic interventions were conducted. This is in line with two recent studies in the field of cardiovascular medicine, one of which investigated sex-based differences in noninvasive diagnostic methods for coronary artery disease in establishing disease in men and women. In this study, women were found to have more normal diagnostic tests than men when presenting symptoms.[26] The other study assessed whether angiographies resulted in sex differences in disease outcome in patients with suspected coronary heart disease. It was found that compared with men, women had a higher likelihood of having no coronary heart disease after undergoing a computed tomography coronary angiography.[27] To the authors' knowledge, no studies assess sex differences in receiving a disease diagnosis after diagnostic interventions were conducted when presenting with common somatic symptoms in general practice.

The current results are also in line with studies that report an increased female prevalence and incidence of functional somatic syndrome diagnoses in general practice.[11] A symptom diagnosis, in which symptoms continue to be symptoms over time, is only to be given if an adequate medical examination and anamnesis has not revealed a condition that explains the symptoms.[28,29] The adequate medical examination comprises diagnostic interventions as discussed here. Thus, if female patients receive relatively more symptom diagnoses, diagnostic interventions relate to fewer disease diagnoses in women.

Clinical Implications

This study shows that differences between women and men are present in the association between diagnostic interventions and disease diagnosis: female patients receive fewer disease diagnoses than male patients after a diagnostic intervention is

conducted. Especially in (lower) back pain, tingling fingers, feet and/or toes, tiredness, and arm and/or leg symptoms, a diagnostic intervention more often yields a disease diagnosis in male patients than in female patients.

At face value, this may point toward a justification of the fewer diagnostic interventions that are performed by the GP in female patients. After all, the odds of receiving a disease diagnosis after an intervention are lower in women than in men. However, recent research also found that the association between female sex and disease diagnosis is mediated by these diagnostic interventions.[3] Although GPs cannot alter whether diagnostic interventions detect disease in women, GPs can increase their rate of diagnostic interventions conducted in women, to avoid underdiagnosis in female patients. Therefore, reducing the use of diagnostic interventions in women even further is unwarranted, as fewer diagnostic interventions in female patients may further exacerbate the difference in disease diagnosis between men and women. Nevertheless, requesting more diagnostic interventions in women should be done with caution, as interventions may cause iatrogenic harm.

Furthermore, as women are thought to seek help earlier in their disease process than men, they may present with more atypical complaints in general practice. These pose a challenge to diagnose, as the symptoms may have not yet progressed enough to be readily attributed to an organic disease, or diagnostic interventions may not be sensitive enough to detect disease. Therefore, further research could focus on sex differences in help-seeking behavior for common somatic symptoms.

In addition, this study could not account for possible underdiagnosis in women. However, it is suggested that the use of sex-specific thresholds for diagnostic interventions, as is done with troponin I, may increase the sensitivity of detecting ischemic heart disease.[30] Thus, further investigations could assess whether diagnostic interventions have a similar diagnostic accuracy in detecting disease or abnormalities in male and female patients presenting with common somatic symptoms.

Unfortunately, our study does not lead to clinical implications that directly result in changes in patient care. The data used in this study do not suffice for that purpose, since a disease diagnosis is not a direct indicator of improved patient care. To formulate a call to action for optimal gender-sensitive patient care in a concrete and evidence-based manner, the current study should be complemented with studies that assess the aforementioned help-seeking for somatic symptoms

and sex-specific diagnostic accuracy for diagnostic interventions in detail. Furthermore, studies that focus on gender-sensitive medicine from the patient's perspective are needed as well. This study functions as a first exploration into possible gender inequities in the primary care process for common somatic symptoms.

References

1. Evangelidou S, NeMoyer A, Cruz-Gonzalez M, O'Malley I, Alegria M. Racial/ethnic differences in general physical symptoms and medically unexplained physical symptoms: Investigating the role of education. *Cultural Diversity & Ethnic Minority Psychology*. 2020;26(4):557-569. doi: 10.1037/cdp0000319.
2. Dirkzwager AJ, Verhaak PF. Patients with persistent medically unexplained symptoms in general practice: Characteristics and quality of care. *BMC Family Practice*. 2007;8(1):33. doi: 10.1186/1471-2296-8-33.
3. Ballering AV, Muijres D, Uijen AA, Rosmalen JGM, olde Hartman TC. Sex differences in the trajectories to diagnosis of patients presenting with common somatic symptoms in primary care: An observational cohort study. *Journal of Psychosomatic Research*. 2021;149:110589. doi: 10.1016/j.jpsychores.2021.110589.
4. Groeneveld JM, Ballering AV, van Boven K, Akkermans RP, Olde Hartman TC, Uijen AA. Sex differences in incidence of respiratory symptoms and management by general practitioners. *Family practice*. 2020;37(5):631-636. doi: 10.1093/fampra/cmaa040.
5. Arber S, McKinlay J, Adams A, Marceau L, Link C, O'Donnell A. Patient characteristics and inequalities in doctors' diagnostic and management strategies relating to CHD: A video-simulation experiment. *Social Science & Medicine*. 2006;62(1):103-115. doi: 10.1016/j.socscimed.2005.05.028.
6. McKinlay JB, Link CL, Freund KM, Marceau LD, O'Donnell AB, Lutfey KL. Sources of variation in physician adherence with clinical guidelines: Results from a factorial experiment. *Journal of General Internal Medicine*. 2007;22(3):289-296.
7. Clerc Liaudat C, Vaucher P, De Francesco T, et al. Sex/gender bias in the management of chest pain in ambulatory care. *Womens Health*. 2018;14:1745506518805641. doi: 10.1177/1745506518805641.
8. Luijks H, van Boven K, olde Hartman T, Uijen A, van Weel C, Schers H. Purposeful incorporation of patient narratives in the medical record in the netherlands. *Journal of the American Board of Family Medicine*. 2021;34(4):709-723. doi: 10.3122/jabfm.2021.04.200609.
9. Zijlema WL, Stolk RP, Löwe B, Rief W, White PD, Rosmalen JGM. How to assess common somatic symptoms in large-scale studies: A systematic review of questionnaires. *Journal of Psychosomatic Research*. 2013;74(6):459-68. doi: 10.1016/j.jpsychores.2013.03.093.
10. Acevedo-Mesa A, Tendeiro JN, Roest A, Rosmalen JGM, Monden R. Improving the measurement of functional somatic symptoms with item response theory. *Assessment*. 2020;28(8):1960-1970. doi: 10.1177/1073191120947153.
11. Verhaak PF, Meijer SA, Visser AP, Wolters G. Persistent presentation of medically unexplained symptoms in general practice. *Family practice*. 2006;23(4):414-420. doi: 10.1093/fampra/cml016.
12. Miles J, Shevlin M. Collinearity. In: *Applying regression and correlation: A guide for students and researchers.* Sage; 2001:126-135.
13. Heidari S, Babor TF, De Castro P, Tort S, Curno M. Sex and gender equity in research: Rationale for the SAGER guidelines and recommended use. *Research Integrity and Peer Review*. 2016;1:2. doi: 10.1186/s41073-016-0007-6.
14. Vinker S, Kvint I, Erez R, Elhayany A, Kahan E. Effect of the characteristics of family physicians on their utilisation of laboratory tests. *The British Journal of General Practice*. 2007;57(538):377-382.
15. Bertakis KD, Azari R. Patient gender and physician practice style. *Journal of Women's Health*. 2007;16(6):859-868.

16. Roter DL, Hall JA, Aoki Y. Physician gender effects in medical communication: A meta-analytic review. *JAMA*. 2002;288(6):756-764. doi: 10.1001/jama.288.6.756.
17. Hall JA, Irish JT, Roter DL, Ehrlich CM, Miller LH. Gender in medical encounters: An analysis of physician and patient communication in a primary care setting. *Health Psychology*. 1994;13(5):384. doi: 10.1037//0278-6133.13.5.384.
18. van Boven K, Uijen AA, van de Wiel N, Oskam SK, Schers HJ, Assendelft WJJ. The diagnostic value of the patient's reason for encounter for diagnosing cancer in primary care. *Journal of the American Board of Family Medicine*. 2017;30(6):806-812. doi: 10.3122/jabfm.2017.06.170076 [doi].
19. Barsky AJ, Peekna HM, Borus JF. Somatic symptom reporting in women and men. *Journal of General Internal Medicine*. 2001;16(4):266-275. doi: 10.1046/j.1525-1497.2001.00229.x.
20. Chakkalakal RJ, Higgins SM, Bernstein LB, et al. Does patient gender impact resident physicians' approach to the cardiac exam? *Journal of General Internal Medicine*. 2013;28(4):561-566. doi: 10.1007/s11606-012-2256-5.
21. Parker R, Larkin T, Cockburn J. A visual analysis of gender bias in contemporary anatomy textbooks. *Social Science & Medicine*. 2017;180:106-113.
22. Michiels-Corsten M, Donner-Banzhoff N. Beyond accuracy: Hidden motives in diagnostic testing. *Family Practice*. 2018;35(2):222-227. doi: 10.1093/fampra/cmx089.
23. van der Weijden T, van Bokhoven MA, Dinant GJ, van Hasselt CM, Grol RP. Understanding laboratory testing in diagnostic uncertainty: A qualitative study in general practice. *Br J Gen Pract*. 2002;52(485):974-980.
24. Lutfey KE, Link CL, Marceau LD, et al. Diagnostic certainty as a source of medical practice variation in coronary heart disease: Results from a cross-national experiment of clinical decision making. *Medical Decision Making*. 2009;29(5):606-618. doi: 10.1177/0272989X09331811.
25. Barsky AJ, Orav EJ, Bates DW. Somatization increases medical utilization and costs independent of psychiatric and medical comorbidity. *Archives of General Psychiatry*. 2005;62(8):903-910. doi: 10.1001/archpsyc.62.8.903.
26. Pagidipati NJ, Coles A, Hemal K, et al. Sex differences in management and outcomes of patients with stable symptoms suggestive of coronary artery disease: Insights from the PROMISE trial. *American Heart Journal*. 2019;208:28-36. doi: 10.1016/j.ahj.2018.11.002.
27. Mangion K, Adamson PD, Williams MC, et al. Sex associations and computed tomography coronary angiography-guided management in patients with stable chest pain. *Eur Heart J*. 2020;41(13):1337-1345. doi: 10.1093/eurheartj/ehz903.
28. Olde Hartman TC, Blankenstein AH, Molenaar AO, et al. NHG-standaard somatisch onvoldoende verklaarde lichamelijke klachten (SOLK). *Huisarts Wet*. 2013;56(5):222-230.
29. Soler JK, Okkes I. Reasons for encounter and symptom diagnoses: A superior description of patients' problems in contrast to medically unexplained symptoms (MUS). *Family practice*. 2012;29(3):272-282. doi: 10.1093/fampra/cmr101.
30. Humphries KH, Izadnegahdar M, Sedlak T, et al. Sex differences in cardiovascular disease - impact on care and outcomes. *Front Neuroendocrinol*. 2017;46:46-70. doi: 10.1016/j.yfrne.2017.04.001.

Appendices

Appendix A. Categorization of comorbidities

Malignancies		**Cardiovascular disease**	
ICPC	**Description**	**ICPC**	**Description**
A79	Malignancy NOS	K72	Cardiovascular neoplasm
B72	Hodgkin's disease/Lymphoma	K74	Ischaemic heart dis w. angina
B73	Leukaemia	K75	Acute myocardial infarction
B74	Malig. neoplasm blood other	K76	Ischaemic heart dis w/o angina
B75	Beningn/unspecified neoplasm blood	K77	Heart failure
D74	Malig. neoplasm stomach	K78	Atrial fibrillation/flutter
D75	Malig. neoplasm colon/rectum	K79	Paroxysmal tachycardia
D76	Malig. neoplasm pancreas	K80	Cardiac arrhythmia NOS
D77	Malig. Neoplasm digest oth/NOS	K81	Heart/arterial murmur NOS
F74	Neoplasm of eye/adnexa	K82	Pulmonary heart disease
H75	Neoplasm of ear	K83	Heart valve disease NOS
L71	Malig. neoplasm musculoskeletal	K84	Heart disease other
N74	Malig. neoplasm nervous system	K86	Hypertension uncomplicated
N75	Benign neoplasm nervous system	K87	Hypertension complicated
N76	Neoplasm nervous system unspec.	K89	Transient cerebral ischaemia
R84	Malig. neoplasm bronchus/lung	K90	Stroke/cerebroovascular accident
R85	Malig. neoplasm respiratory, other	K91	Cerebrovascular disease
R92	Neoplasm respiratory unspecified	K92	Atherosclerosis/PVD
T71	Malig. neoplasm thyroid		
T73	Neoplasm endocrine oth/unspecified	**Asthma and COPD**	
U75	Malig. neoplasm of kidney	**ICPC**	**Description**
U76	Malig. neoplasm of bladder	R95	Chronic obstructive pulmonary dis
U77	Malig. neoplasm urinary other	R96	Asthma
U79	Neoplasm urinary tract NOS	*Abbreviations: "dis" disease, "malig." malignant, "NOS" not otherwise specified, "oth "other, "unspec." Unspecified, "w" with, "w/o" without.	
X75	Malig. neoplasm cervix		
X76	Malig. neoplasm breast female		
X77	Malig. Neoplasm female genital other		
X81	Genital neoplasm female oth/unspec.		
Y77	Malig. neoplasm prostate		
Y78	Malig. neoplasm male genital other		

Appendix B. Frequency of performed diagnostic interventions, stratified by sex and RFE.

Interventions	Number of EoC (%)			
	Female	Male	Total	*p*-value[a]
Interventions for headache (N01; N=3,411)				
Physical examination	1,615 (75.0)	977 (77.6)	2,592 (76.0)	0.09
Laboratory diagnostics	255 (11.8)	154 (12.2)	409 (12.0)	0.74
Diagnostic imaging	30 (1.4)	21 (1.7)	51 (1.5)	0.53
Referral to specialist	223 (10.4)	141 (11.2)	364 (10.7)	0.45
Interventions for dizziness (N17; N=2,709)				
Physical examination	1,461 (82.0)	778 (83.8)	2,239 (82.7)	0.24
Laboratory diagnostics	558 (31.3)	201 (21.7)	759 (28.0)	<0.001
Diagnostic imaging	10 (0.6)	9 (1.0)	19 (0.7)	0.23
Referral to specialist	170 (9.5)	115 (12.4)	285 (10.5)	0.22
Interventions for heart pain (K01; N=660)				
Physical examination	261 (77.4)	261 (80.8)	522 (79.1)	0.29
Laboratory diagnostics	73 (21.7)	66 (20.4)	660 (21.1)	0.70
Diagnostic imaging	6 (1.8)	7 (2.2)	13 (2.0)	0.72
Referral to specialist	114 (33.8)	134 (41.5)	248 (37.6)	0.04
Interventions for (lower) back pain (L02/L03; N=5,962)				
Physical examination	2,620 (75.2)	1,938 (78.2)	4,558 (76.5)	<0.01
Laboratory diagnostics	359 (10.3)	153 (6.2)	512 (8.6)	<0.001
Diagnostic imaging	502 (14.4)	339 (13.7)	841 (14.1)	0.43
Referral to specialist	244 (7.0)	180 (7.3)	424 (7.1)	0.69
Interventions for nausea (D09; N=1,222)				
Physical examination	568 (63.0)	223 (69.5)	791 (64.7)	0.04
Laboratory diagnostics	203 (22.5)	69 (21.5)	272 (22.2)	0.70
Diagnostic imaging	34 (3.8)	10 (3.1)	44 (3.6)	0.59
Referral to specialist	86 (9.5)	45 (14.0)	131 (10.7)	0.03
Interventions for muscle pain (L18; N=554)				
Physical examination	219 (68.0)	171 (73.7)	390 (70.4)	0.15
Laboratory diagnostics	98 (30.4)	79 (34.1)	177 (31.9)	0.37
Diagnostic imaging	11 (3.4)	13 (5.6)	24 (4.3)	0.21
Referral to specialist	36 (11.2)	22 (9.5)	58 (10.5)	0.52
Interventions for shortness of breath (R02; N=3,593)				
Physical examination	1,903 (90.6)	1,375 (92.1)	3,278 (91.2)	0.12
Laboratory diagnostics	559 (26.6)	376 (25.2)	935 (26.0)	0.33
Diagnostic imaging	159 (7.6)	151 (10.1)	310 (8.6)	<0.01
Referral to specialist	302 (14.4)	263 (17.6)	565 (15.7)	<0.01

Appendix B. Continued

Interventions	Number of EoC (%)			
	Female	Male	Total	*p*-value[a]
Interventions for chills (A02; N=132)				
Physical examination	39 (58.2)	54 (83.1)	93 (70.5)	<0.01
Laboratory diagnostics	20 (29.9)	16 (24.6)	36 (27.3)	0.50
Diagnostic imaging	1 (1.4%)	1 (1.5%)	2 (1.5%)	0.98
Referral to specialist	10 (14.9)	14 (21.5)	24 (18.2)	0.33
Interventions for tingling in fingers, feet and/or toes (N05; N=609)				
Physical examination	311 (85.2)	204 (83.6)	515 (84.6)	0.59
Laboratory diagnostics	72 (19.7)	55 (22.5)	127 (20.9)	0.40
Diagnostic imaging	3 (0.8)	6 (2.5)	9 (1.5)	0.10
Referral to specialist	61 (16.7)	33 (13.5)	94 (15.4)	0.29
Interventions for throat symptoms (D21/R21; N=4,961)				
Physical examination	2,498 (82.3)	1,631 (84.6)	4,129 (83.2)	0.03
Laboratory diagnostics	383 (12.6)	211 (10.9)	594 (12.0)	0.08
Diagnostic imaging	36 (1.2)	25 (1.3)	61 (1.2)	0.73
Referral to specialist	243 (8.0)	165 (8.6)	408 (8.2)	0.49
Interventions for weakness (A04; N=5,608)				
Physical examination	2,599 (70.4)	1,467 (76.6)	4,066 (72.5)	<0.001
Laboratory diagnostics	2,759 (74.7)	1,311 (68.5)	4,070 (72.6)	<0.001
Diagnostic imaging	104 (2.8)	73 (3.8)	177 (3.2)	0.04
Referral to specialist	284 (7.7)	200 (10.4)	484 (8.6)	<0.001
Interventions for arm and leg symptoms (L09/L14; N=4,847)				
Physical examination	2,408 (86.4)	1,797 (87.2)	4,205 (86.8)	0.40
Laboratory diagnostics	310 (11.1)	184 (8.9)	494 (10.2)	<0.01
Diagnostic imaging	368 (13.2)	296 (14.4)	552 (11.4)	0.24
Referral to specialist	387 (13.9)	299 (14.5)	686 (14.2)	0.54
Interventions for all RFE (N=34,268)				
Physical examination	16,502 (78.5)	10,876 (82.1)	27,378 (79.9)	<0.001
Laboratory diagnostics	5,649 (26.9)	2,875 (21.7)	8,524 (24.9)	<0.001
Diagnostic imaging	1,264 (6.0)	951 (7.2)	2,215 (6.5)	<0.001
Referral to specialist	2,160 (10.3)	1,611 (12.2)	3,771 (11.0)	<0.001

[a]chi square test

Appendix C. Significant diagnostic intervention-by-female sex interaction terms

Independent variables	OR (95% CI)[a]			
	Headache (N=3,411)	**Dizziness (N=2,709)**	**Heart pain (N=660)**	**(Lower) back pain (N=5,962)**
Physical examination	ns	ns	ns	0.51 (0.34-0.78)
Laboratory diagnostics	ns	ns	ns	ns
Imaging	ns	ns	ns	0.64 (0.43-0.96)
Specialist referral	ns	ns	0.39 (0.17-0.90)	0.46 (0.29-0.72)
	OR (95% CI)[a]			
	Nausea (N=1,222)	**Muscle pain (N=554)**	**Shortness of breath (N=3,593)**	**Chills (N=132)**
Physical examination	ns	ns	ns	ns
Laboratory diagnostics	ns	ns	ns	ns
Imaging	ns	ns	ns	ns
Specialist referral	ns	ns	ns	ns
	OR (95% CI)[a]			
	Tingling in fingers, feet and/or toes (N=609)	**Swallowing/ throat symptoms (N=4,961)**	**Weakness/ general tiredness (N=5,608)**	**Arm and/ or leg symptoms (N=4,847)**
Physical examination	ns	ns	ns	ns
Laboratory diagnostics	0.32 (0.13-0.81)	ns	0.72 (0.53-0.96)	0.56 (0.37-0.86)
Imaging	ns	ns	ns	ns
Specialist referral	ns	ns	ns	0.65 (0.46-0.92)

[a]Adjusted for the patients' age at diagnosis, comorbidities, number of contacts per EoC and sex.

CHAPTER 11

Ballering, A.V., Oertelt-Prigione, S., Lifelines Corona Research Initiative, olde Hartman, T.C., & Rosmalen, J.G.M. (2021). Sex and gender-related differences in COVID-19 diagnoses and SARS-CoV-2 testing practices during the first wave of the pandemic: The Dutch lifelines COVID-19 cohort study. *Journal of Women's Health*, *30*(12), 1686-1692.

Abstract

Background: Although sex differences are described in COVID-19 diagnoses and testing, many studies neglect possible gender-related influences. Additionally, research is often performed in clinical populations, while most COVID-19 patients are not hospitalized. Therefore, we investigated associations between sex and gender-related variables, and COVID-19 diagnoses and testing practices in a large general population cohort during the first wave of the pandemic when testing capacity was limited.

Methods: We used data from the Lifelines COVID-19 Cohort (N=74,722; 60.8% female). We applied bivariate and multiple logistic regression analyses. The outcomes were a COVID-19 diagnosis (confirmed by SARS-CoV-2 PCR testing or physician's clinical diagnosis) and PCR testing. Independent variables included among others participants' sex, age, somatic comorbidities, occupation and smoking status. Sex-by-comorbidity and sex-by-occupation interaction terms were included to investigate sex differences in associations between the presence of comorbidities or an occupation with COVID-19 diagnoses or testing practices.

Results: In bivariate analyses female sex was significantly associated with COVID-19 diagnoses and testing, but significance did not persist in multiple logistic regression analyses. However, a gender-related variable, being a healthcare worker, was significantly associated with COVID-19 diagnoses (OR=1.68; 95%CI=1.30-2.17) and testing (OR=12.5; 95%CI=8.55-18.3). Female healthcare workers were less often diagnosed and tested than male healthcare workers ($OR_{interaction}$=0.54; 95%CI=0.32-0.92, $OR_{interaction}$=0.53; 95%CI=0.29-0.97, respectively).

Conclusion: We found no sex differences in COVID-19 diagnoses and testing in the general population. Among healthcare workers a male preponderance in COVID-19 diagnoses and testing was observed. This could be explained by more pronounced COVID-19 symptoms in males or by institutionalized gender inequities.

Introduction

Globally, males and females are infected with Severe Acute Respiratory Syndrome-Coronavirus-2 (SARS-CoV-2), the virus responsible for COVID-19, at equal rates.[1,2] However, sex differences are found in multiple other aspects of COVID-19. Males experience higher rates of hospitalizations, intensive care unit admission and COVID-19-related deaths.[1,3] These differences can be partly explained by biological sex differences, for example in innate and humoral immune responses[4] or in rates of pre-existing somatic comorbidities, such as cardiovascular disease, which associate with a poor COVID-19 prognosis.[5-7]

In addition to these sex-specific differences, gender-related factors associate with the course of SARS-CoV-2 infection. Gender is the embodiment of different roles, behaviors, identities and relationships by individuals according to societal norms, which result in different expectations, opportunities and experiences.[8] It can modify risk factor distribution and exposure patterns.[3] For example, women constitute the majority of the health workforce worldwide, and are more likely to work in the service industry and contact professions.[9] A recent study including 99,795 healthcare workers and 2,035,395 community individuals showed that healthcare workers had a 3.4-fold greater risk of infection with SARS-CoV-2, compared to the general population.[10] Epidemiological data about COVID-19 are thus influenced by sex differences and gendered health-related behaviors and roles. Disentangling the contribution of these factors is complex as the presence of multilayered interactions have to be assumed.[11]

Gender can also affect access to and uptake of diagnostic measures. A Canadian study including 233,566 individuals that received SARS-CoV-2 tests, demonstrated that 64.3% of the conducted tests were performed in women.[12] Two other Canadian general population studies including 409,207 and 4,240 individuals showed no difference or a small skewing towards the uptake of tests in women, respectively.[13]

Most medical research on sex differences in COVID-19 is conducted in clinical populations, which represent a fraction of the affected population. Hence, we are currently unable to formulate healthcare system-wide implications and recommendations based on these data.[14] Additionally, most clinical research and large-scale epidemiological research neglects the possible influence of gender. Few previous studies accounted for participants' occupation and the gendered aspect hereof. Given the significant over-representation of women in the healthcare workforce and other contact professions, such as those in education, it is important to adjust for these factors when assessing sex and gender-related differences in health outcomes.

Both sex and gender-related risk factors are known to unequally affect men's and women's health and access to healthcare.[15,16] First, obtaining insights into possible influences of sex and gender during the first wave of the COVID-19 pandemic is pivotal for understanding health-related inequities between women and men during times of health-resource scarcity. Secondly, these insights may bear important policy implications for future disaster preparedness. Thirdly, they might be especially relevant for low- and middle income countries where health-resource scarcity during an epidemic or pandemic may persist longer due to pre-existing resource constraints. Lastly, to inform public health policies, conclusions from clinical research studies do not suffice and general population studies should be performed as well.

Therefore, we investigated the associations between sex and gender-related factors with COVID-19 diagnoses and testing practices in the Lifelines COVID-19 Cohort study, which includes 74,722 unique participants from the general population of the North of the Netherlands. We hypothesize that both female sex and feminine gendered factors associate with a COVID-19 diagnosis and SARS-CoV-2 testing practices.

Methods

Data source

This study is based on data collected within the Dutch Lifelines Cohort Study and its digital add-on study, the Lifelines COVID-19 cohort. The former is a multi-disciplinary prospective population-based cohort study examining in a unique three-generation design the health and health-related behaviors of 167,729 persons living in the North of The Netherlands. It employs a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioral, physical and psychological factors which contribute to the health and disease of the general population, with a special focus on multi-morbidity and complex genetics.

We included data of 13 consecutive Lifelines COVID-19 measurements from March 2020 to August 2020. Initially questionnaires for continuous data collection were send out weekly, from June 2020 onwards data was collected at biweekly intervals. Participants of the Lifelines COVID-19 cohort are recruited from the Lifelines population cohort and the Lifelines NEXT birth cohort. The Lifelines Cohort Study is performed according to the principles of the Declaration of Helsinki. The Lifelines Cohort Study is approved by the Medical Ethical Committee of the University Medical Center Groningen (number: 2007/152). All participants provided written consent. Extensive information on the cohorts, design considerations and recruitments procedures is provided elsewhere.[17-20]

COVID-19 and the testing regime in the North of the Netherlands

The first confirmed COVID-19 case in the Netherlands was reported on 27 February 2020. A month later the number of new cases diagnosed per day reported by the municipal health services was 1,178. The number of deaths related to COVID-19 had risen as well, with a peak of 178 deaths per day on 2 April 2020[21]. The steep rise in cases and case-fatalities prompted the Dutch government to announce a nation-wide lockdown on 15 March 2020.

In the Northern provinces of the Netherlands the COVID-19 outbreak followed a different pattern than in other regions of the country. The first COVID-19 cases were reported on 1 March 2020, 10 March 2020 and 11 March 2020 in the Northern provinces of Drenthe, Friesland and Groningen, respectively.[20] Up until 9 June 2020, merely 3.1%, 2.7% and 2.0% of the national cumulative SARS-CoV-2 infections, hospitalizations and COVID-19-related deaths were reported in the North of the Netherlands, respectively. Additional information on the COVID-19 outbreak, its facilitators and barriers in the Northern Dutch provinces can be found elsewhere.[20]

Notably, the testing regime in the Netherlands was restricted during the first wave of the COVID-19 pandemic, as a shortage of reagents occurred during the first months of the pandemic.[22] This meant that healthcare workers and severely ill patients with a suspected SARS-CoV-2 infection were prioritized in testing procedures. The limited availability of SARS-CoV-2 testing equipment for the general population persisted until August 2020.

Participants

Participants over 18 years of age completed digital questionnaires on multiple topics, including but not limited to demographics, occupation, physical and mental health, and adherence to COVID-19 guidelines. The questionnaires register the participants' municipally-registered sex, which generally corresponds to biological sex-at-birth, hence we refer to the participants as male and female. The questionnaire does not include a specific question about gender identity. Of the 74,722 unique participants (aged 18 to 94 years) included in this study, 61,584 (82.4%) unique participants were included during the first three measurements (***Appendix A***). In total, 30,326 (40.6%) participants completed ten or more questionnaires. Additional information about the study population is shown in ***Table 1.***We did not find any indication for relevant systemic attrition: no meaningful associations between potential predictors of infection or testing and attrition rates existed. Similarly, proportions of missing data did not differ meaningfully between groups (e.g. male/female or infected/non-infected).

Variables and statistical analyses

Descriptive statistics are provided as absolute numbers with concomitant percentages. If appropriate, means with standard deviations are provided. Data were examined for normality using q-q plots and histograms.

The outcome variables included participants' receipt of a COVID-19 diagnosis and test for SARS-CoV-2. We defined a COVID-19 diagnosis as either a self-reported positive SARS-CoV-2 PCR test or a self-reported clinical diagnosis of COVID-19 by a physician based on participants' symptoms. We provide an overview of the population's characteristics stratified by diagnosis method (***Appendix B***).

All independent variables were self-reported and the entry given during the first completed questionnaire was included in analyses. The presence of chronic diseases covered a range of somatic diseases, including but not limited to cardiovascular, autoimmune and neurological diseases (***Appendix C***). Age and biological sex were derived from the municipal registration database. Participants' educational level was derived from the Lifelines Cohort Study data and defined as described earlier.[23] Participants' type of contact profession was derived from both the Lifelines Cohort Study data and the Lifelines COVID-19 Cohort. The former provided data, up until 2018, on whether participants' were healthcare or educational professionals (ISCO sub-major group 22 and 23, respectively), while the COVID-19 Cohort provided information about whether participants had a contact profession.

To identify whether sex and gender-related variables associate with COVID-19 diagnoses, we conducted multiple logistic regression analyses. Participants' sex, age in years, educational level, presence of chronic somatic diseases, smoking status, adherence to mitigation guidelines, household composition, working place and (contact requiring) occupation were included as independent variables. We included interaction terms between sex and the presence of a chronic disease, and between sex and occupation to assess whether the association between the respective independent variables and outcome differed per sex. Participants with missing data on independent variables were excluded listwise.

We conducted a similar multiple logistic regression analysis with SARS-CoV-2 testing as an outcome, but with fewer independent variables due to fewer testing events. Participants' biological sex, age in years, educational attainment, presence of chronic somatic disease, smoking status, household composition, working place and (contact requiring) occupation were included as independent variables. Interaction terms between sex and the presence of chronic disease, and sex and occupation were included in the analyses as well.

To assess whether participants' age in years fulfilled the linearity assumption of multiple logistic regression analyses, we divided the variable into quartiles and assessed whether the odds of being diagnosed or tested were monotonically changing. We maintained a two-sided α-value, corrected for multiple comparisons, of 0.002 (0.05/23, 19 predictors and 4 sex-by-variable interaction terms within a family of tests). IBM SPSS v. 25 was used to perform analyses.

Results

In total, 544,077 questionnaires were completed by 74,722 unique participants (60.8% female). ***Table 1*** provides an overview of the characteristics of the study population.

COVID-19 diagnosis

In bivariate analyses female sex associated with a COVID-19 diagnosis: females had 1.37 (95%CI=1.21-1.55) times the odds of males to be diagnosed with COVID-19 (***Appendix D***). However, ***Table 2*** shows that this association did not persist if adjusted for working a contact profession and additional variables (OR=0.94; 95%CI=0.81-1.09). Working in a healthcare profession was found to be positively associated with a COVID-19 diagnosis, (OR=1.68; 95%CI=1.30-2.17). In female participants the association between being a healthcare worker and a COVID-19 diagnosis was statistically different (OR=1.84; 95%CI=1.61-2.12) from that in male participants (OR=2.69; 95%=1.66-4.38). The interaction term show female healthcare workers were diagnosed less often than their male counterparts ($OR_{interaction}$=0.54; 0.32-0.92).

SARS-CoV-2 testing

The bivariate analyses in ***Appendix D*** show that female sex was statistically significantly associated with SARS-CoV-2 PCR testing (OR=2.04; 95%CI=1.74-2.41). However, as shown by ***Table 3***, female sex was not significantly associated with SARS-CoV-2 testing upon adjustment for additional variables (OR=1.04; 95%CI=0.77-1.42). Healthcare workers had higher odds of being tested for SARS-CoV-2 (OR=12.5; 95%CI=8.55-18.3) than the reference category. Notably, female healthcare workers were less often tested than their male counterparts ($OR_{interaction}$=0.53; 95%CI=0.29-0.97).

Table 1. Characteristics of the study population

		Male participants (N=29,273; 39.2%)	Female participants (N=45,449; 60.8%)
Age in years, mean (SD)		55.4 (12.9)	52.7 (12.9)
Educational attainment, N (%)	Low	4,209 (14.4)	5,113 (11.2)
	Medium	13,351 (45.6)	23,820 (52.4)
	High	11,319 (38.7)	15,768 (34.7)
PCR test conducted, N (%)		191 (0.7)	602 (1.3)
Positive PCR test, N (%)		52 (0.2)	131 (0.3)
Positive doctor's diagnosis, N (%)		322 (1.1)	661 (1.5)
COVID-19 diagnosis[a], N (%)		355 (1.2)	751 (1.7)
Hospitalized for a COVID-19 diagnosis, N (%)		13 (0.0)	<10 (0.0)
Chronic disease, N (%)	Yes	5,896 (20.1)	10,526 (23.2)
	No	16,409 (56.1)	24,469 (53.8)
Smoking, N (%)	Yes	2,896 (9.9)	3,979 (8.8)
	No	25,825 (88.2)	40,634 (89.4)
Alcohol, N (%)	Yes	19,602 (67.0)	23,010 (50.6)
	No	9,089 (31.0)	21,562 (47.4)
Precautions taken, N (%)	Frequently washing hands or use of desinfectant	27,293 (93.2)	43,542 (95.8)
	Social distancing	28,095 (96.0)	43,831 (96.4)
	Avoid use of public transport	20,092 (68.6)	33,474 (73.7)
	Covering nose and mouth in public	2,770 (9.5)	4,126 (9.1)
Household includes a child, N (%)		7,669 (26.2)	12,145 (26.7)
Household includes another adult, N (%)		15,547 (53.1)	22,780 (50.1)
Household includes an elderly person, N (%)		7,715 (26.4)	13,045 (28.7)
Profession that requires contact, N (%)	No	15,416 (52.7)	15,782 (34.7)
	Yes, but no education or health professional	2,132 (7.3)	7,062 (15.5)
	Yes, as an education professional	1,418 (4.8)	3,454 (7.6)
	Yes, as a healthcare professional	670 (2.3)	3,034 (6.7)
	Not reported	9,562 (32.9)	16,146 (35.5)

[a]Defined as receiving either a positive SARS-CoV-2 PCR test, or a positive clinician's diagnosis

Table 2. Associations between predictors and COVID-19 diagnosis.

		Total population (N=74,722) OR (95% CI)	Male participants (N=29,273) OR (95% CI)	Female participants (N=45,449) OR (95% CI)
Female sex		0.94 (0.81-1.09)	n.a.	n.a.
Age		0.99 (0.99-1.00)	1.01 (0.99-1.02)	0.99 (0.98-1.00)
Educational attainment	Low	1.00 (ref)	1.00 (ref)	1.00 (ref)
	Medium	0.94 (0.72-1.23)	0.78 (0.53-1.16)	1.06 (0.74-1.53)
	High	0.68 (0.52-0.91)	0.69 (0.45-1.05)	0.69 (0.47-1.03)
Chronic disease present		**1.34 (1.15-1.55)[a]**	1.18 (0.91-1.52)	**1.48 (1.36-1.62)**
Smoking		0.92 (0.71-1.19)	0.86 (0.57-1.32)	0.89 (0.76-1.04)
Frequent handwashing and use of desinfectant		0.72 (0.45-1.15)	0.76 (0.42-1.38)	0.69 (0.32-1.47)
Social distancing		**0.22 (0.11-0.44)**	**0.26 (0.09-0.79)**	**0.21 (0.13-0.35)**
Avoidance of public transport		1.16 (0.98-1.37)	1.14 (0.87-1.49)	1.01 (0.92-1.11)
Covering nose and mouth in public		**1.61 (1.33-1.97)**	**1.53 (1.10-2.13)**	**1.99 (1.78-2.22)**
Household members ≤ 18 years		0.98 (0.83-1.16)	0.96 (0.73-1.27)	1.00 (0.91-1.09)
Household members 19-59 years		**0.51 (0.42-0.60)**	**0.43 (0.32-0.59)**	**0.79 (0.71-0.89)**
Household members ≥ 60 years		**0.76 (0.61-0.94)**	**0.63 (0.42-0.93)**	0.91 (0.79-1.03)
Working from home		**1.51 (1.30-1.78)**	**1.68 (1.28-2.19)**	1.02 (0.93-1.13)
Contact profession	No	1.00 (ref)	1.00 (ref)	1.00 (ref)
	Yes, but not in education or healthcare	**1.22 (1.02-1.47)[b]**	**1.43 (1.04-1.98)**	**1.34 (1.12-1.49)**
	Yes, in education	**1.38 (1.08-1.77)[c]**	1.07 (0.65-1.75)	**1.32 (1.13-1.53)**
	Yes, in healthcare	**1.68 (1.30-2.17)[d]**	**2.69 (1.66-4.38)**	**1.84 (1.61-2.12)**

[a] Sex-by-chronic disease interaction: OR=1.17, 95%CI=0.86-1.60; [b] Sex-by-contact profession interaction: OR=0.91, 95%CI=0.64-1.31; [c] Sex-by-education profession interaction: OR=1.29, 95%CI=0.75-2.20; [d] Sex-by-healthcare profession interaction: OR=0.54, 95%CI=0.32-0.92. Bold values indicate significance below $p<0.002$.

Table 3. Associations between predictors and receiving SARS-CoV-2 PCR tests.

		Total population (N=74,722) OR (95%CI)	Male participants (N=29,273) OR (95% CI)	Female participants (N=45,449) OR (95% CI)
Female sex		1.04 (0.77-1.42)	n.a.	n.a.
Age		**0.97 (0.95-0.98)**	0.97 (0.95-1.00)	**0.96 (0.95-0.98)**
Educational attainment	Low	1.00 (ref)	1.00 (ref)	1.00 (ref)
	Medium	1.10 (0.63-1.93)	0.50 (0.19-1.30)	1.39 (0.67-2.89)
	High	1.18 (0.66-2.10)	0.79 (0.31-2.04)	1.31 (0.62-2.79)
Chronic disease present		0.94 (0.70-1.26)[a]	0.85 (0.45-1.59)	0.96 (0.69-1.33)
Smoking		0.67 (0.39-1.16)	0.19 (0.03-1.35)	0.84 (0.47-1.49)
Household members ≤ 18 years		1.03 (0.78-1.36)	0.98 (0.54-1.80)	1.07 (0.78-1.47)
Household members 19-59 years		**0.57 (0.41-0.80)**	**0.46 (0.22-0.95)**	**0.60 (0.41-0.87)**
Household members ≥ 60 years		0.80 (0.53-1.21)	1.09 (0.45-2.62)	0.72 (0.45-1.16)
Working from home		**0.46 (0.33-0.66)[b]**	1.13 (0.62-2.06)	**0.30 (0.18-0.47)**
Contact profession	No	1.00 (ref)	1.00 (ref)	1.00 (ref)
	Yes, but not in education or healthcare	**2.86 (1.96-4.17)[c]**	**2.70 (1.26-5.80)**	**2.52 (1.59-3.99)**
	Yes, in education	1.41 (0.70-2.83)[d]	1.89 (0.63-5.65)	1.16 (0.47-2.86)
	Yes, in healthcare	**12.5 (8.55-18.3)[e]**	**21.4 (11.0-41.6)**	**10.1 (6.34-16.1)**

[a] Sex-by-chronic disease interaction: OR=1.04, 95%CI=0.52-2.09; [b] Sex-by-working from home interaction: OR=0.27, 95%CI=0.13-0.53; [c] Sex-by-contact profession interaction: OR=2.28, 95%CI=1.10-4.72; [d] Sex-by-educational profession interaction: OR=0.38, 95%CI=0.10-1.42; [e] Sex-by-healthcare profession interaction: OR=0.53, 95%CI=0.29-0.97. Bold values indicate significance below $p<0.002$. n.a., not applicable

Discussion

This study is the first large epidemiological study in the general population that assesses sex and gender-related differences in COVID-19 diagnosis and SARS-CoV-2 testing in the first wave of the COVID-19 pandemic in the Netherlands. In bivariate analyses we found that females had higher odds than males on receiving a SARS-CoV-2 PCR and a COVID-19 diagnosis. However, upon adjustment for additional covariates, these sex differences did not persist, which is in contrast with our hypothesis. Healthcare workers received significantly more tests and COVID-19 diagnoses than other employee groups, with male healthcare workers receiving significantly more tests and COVID-19 diagnoses than female healthcare workers.

Strengths and limitations

The principal strength of this study was that the Dutch Lifelines Cohort is a large and already established cohort study. Therefore, information about participants was registered before the COVID-19 pandemic. This allowed us to include more information in our analyses than the Lifelines COVID-19 questionnaires. Another strength is that data was collected during the first wave of the COVID-19 pandemic, during which a scarcity in SARS-CoV-2 PCR tests occurred, allowing for more pronounced appearance of potential biases.

However, our study was limited by the relatively few COVID-19 cases in the North of the Netherlands. Moreover, we considered participants with a self-reported physician's COVID-19 diagnosis based on symptoms as COVID-19 positive, while a recent study shows that symptoms only moderately associate with a positive SARS-CoV-2 PCR test.[24] Therefore, as physicians diagnosed patients with COVID-19 based on their symptoms, we may have overestimated the proportion of COVID-19 positive participants (i.e. participants with detectable SARS-CoV-2) in our study population. We also included one's profession as a gender-related variable, but we did not account for other gender-related variables, for example unpaid (child) care responsibilities.

Furthermore, healthcare workers also differ in their professional tasks and disciplinary expertise, from intensivists to home nurses for example, thus displaying a range of occupational risk for COVID-19. Similarly, healthcare workers may have different hierarchical positions in the hospital, which may associate to the probability of a participant's uptake of a SARS-CoV-2 test. Additionally, participants may have changed occupation or retired between completing the occupational questionnaire up to 2018 in the Dutch Lifelines Cohort and the Lifelines COVID-19 cohort questionnaires. Lastly, as the presence and severity of COVID-19-related symptoms may be worse in men, possibly due to a less adequate innate and humoral immune response,[25] this may have introduced a male bias in the access to SARS-CoV-2 testing as this often depends on the presence of symptoms.

Comparison to literature

Although bivariate analyses showed that females in the general population were more often diagnosed with COVID-19 than males, these sex differences did not persist in multiple regression analyses. This is in line with previous research, including a recent meta-analysis of 3,111,714 cases, demonstrating that females and males have confirmed COVID-19 diagnoses at equal rates.[1,25,26]

The association between being a healthcare worker and receiving a COVID-19 diagnosis, in contrast, differed significantly between females and males: male healthcare workers had higher odds of a COVID-19 diagnosis than female counterparts. This seems to contradict earlier Dutch and Canadian research that focused on confirmed and suspected SARS-CoV-2 infection among healthcare workers. In these studies 83.3% and 81.7% of the COVID-19 cases in healthcare workers were female, respectively.[27,28] However, these studies did not adjust for the overrepresentation of females in the population of healthcare workers.

In contrast, a large study including 10,034 healthcare workers from the United Kingdom showed that male staff had an increased risk (OR=1.19; 95%CI=1.01-1.40) of SARS-CoV-2 IgG seropositivity, and likely had been infected with SARS-CoV-2, compared to female staff.[29] Our finding also explains the apparent discrepancy between the observed equal infection rates in men and women,[1] and the increased risk of healthcare workers, of which the majority are women[10]: although the odds of a COVID-19 diagnosis are increased in healthcare workers, especially male healthcare workers appear to be at risk.

In initial bivariate analyses we found that females had higher odds than males for SARS-CoV-2 testing. This is in line with descriptive Canadian studies that assessed sex-disaggregated data on PCR testing.[12,13] However, upon adjustment for additional variables, such as occupation, the identified sex difference was no longer statistically significant.

In healthcare workers this pattern was fully revised with male healthcare workers being at significantly higher odds of SARS-CoV-2 testing than female healthcare workers. Several studies argue that sex- and gender-related factors associate with male-biased patterns of access to diagnostic measures,[30,31] which is in line with our study. We identified more testing in male than female healthcare workers, which might be related to symptoms being possibly more pronounced in males than in females.[32,33] Moreover, gendered work segregation might also play an important role. In fact, a larger proportion of male healthcare workers are employed by hospitals, and not in domestic care or elderly care, where access to personal protective equipment and testing equipment was limited potentially resulting in not only undertesting, but also underdiagnosis.[34]

Implications for Practice and Policy

Conclusively, bivariate analyses demonstrate sex differences in COVID-19 diagnoses and SARS-CoV-2 testing during the first wave of the pandemic in the general population, but these do not remain upon adjustment for additional covariates.

Although a male preponderance in COVID-19 diagnoses and SARS-CoV-2 testing was found among healthcare workers, this may be related to the more pronounced COVID-19 symptoms and worse prognosis in males, respectively. Gender-related factors, such as occupation were found to be relevant in COVID-19 diagnoses and SARS-CoV-2 testing.

Therefore, further research could focus on the independent effects of sex and gender-related factors, besides occupation, on uptake of SARS-CoV-2 tests and COVID-19 diagnoses. For example, it should be investigated in detail whether the increased male test uptake among healthcare workers is due to more males working in hospital settings, in which more testing equipment was available than in female-dominated extramural and domestic care settings.

Moreover, the situation experienced during the first wave of the pandemic in the Netherlands may mimic a more persistent situation in lower and middle income countries. Therefore, research in lower and middle income countries could assess whether a comparable pattern (i.e. predominant testing in male healthcare workers) exists, which may inform policy-making decisions on health-related gender equality.

Additionally, research should focus on whether a sex/gender-bias remains once an open-to-all testing policy is in place, as this would point towards differences in willingness to test, instead of towards limitations in availability of tests. For policymakers, this study implies that conclusions drawn on sex and gender biases in COVID-19 diagnoses and testing from clinical studies should be interpreted with caution, as these cannot be readily extrapolated to other (care) settings, such as the general population or primary and elderly care. Overall, the identified bias in testing procedures should be evaluated on a larger scale to assess its justification or remove its inherent bias.

References

1. Global Health-50/50. The COVID-19 sex-disaggregated data tracker: November update report. https://globalhealth5050.org/the-sex-gender-and-covid-19-project/. Updated 2020. Accessed 16 December, 2020.
2. Salvati L, Biagioni B, Vivarelli E, Parronchi P. A gendered magnifying glass on COVID-19. *Clinical and molecular allergy CMA*. 2020;18(1):14. doi: 10.1186/s12948-020-00129-2.
3. Ahmed SB, Dumanski SM. Sex, gender and COVID-19: A call to action. *Canadian journal of public health*. 2020;111(6):980-983.
4. Scully EP, Haverfield J, Ursin RL, Tannenbaum C, Klein SL. Considering how biological sex impacts immune responses and COVID-19 outcomes. *Nature reviews. Immunology*. 2020;20(7):442-447.
5. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in wuhan, china: A retrospective cohort study. *The Lancet*. 2020;395(10229):1054-1062. doi: 10.1016/S0140-6736(20)30566-3.
6. Grasselli G, Zangrillo A, Zanella A, et al. Baseline characteristics and outcomes of 1591 patients infected with SARS-CoV-2 admitted to ICUs of the lombardy region, italy. *JAMA*. 2020;323(16):1574-1581.
7. Petrilli CM, Jones SA, Yang J, et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in new york city: Prospective cohort study. *BMJ*. 2020;369:m1966. doi: 10.1136/bmj.m1966.
8. Johnson JL, Greaves L, Repta R. Better science with sex and gender: Facilitating the use of a sex and gender-based analysis in health research. *International Journal for Equity in Health*. 2009;8(1):14. doi: 10.1186/1475-9276-8-14.
9. International Labour Office. Women at work trends 2016. https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_457317.pdf. Updated 2016. Accessed December 1, 2020.
10. Nguyen LH, Drew DA, Graham MS, et al. Risk of COVID-19 among front-line health-care workers and the general community: A prospective cohort study. *The Lancet Public health*. 2020;5(9):e475-e483. doi: 10.1016/S2468-2667(20)30164-X.
11. Krieger N, Chen JT, Waterman PD. Excess mortality in men and women in massachusetts during the COVID-19 pandemic. *The Lancet*. 2020;395(10240):1829. doi: 10.1016/S0140-6736(20)31234-4.
12. Stall NM, Wu W, Lapointe-Shaw L, et al. Sex- and Age-Specific differences in COVID-19 testing, cases, and outcomes: A Population-Wide study in ontario, canada. *Journal of the American Geriatrics Society*. 2020;68(10):2188-2191. doi: 10.1111/jgs.16761.
13. Lapointe-Shaw L, Rader B, Astley CM, et al. Web and phone-based COVID-19 syndromic surveillance in canada: A cross-sectional study. *PLoS ONE*. 2020;15(10):e0239886. doi: 10.1371/journal.pone.0239886.
14. Green LA, Fryer GE, Yawn BP, Lanier D, Dovey SM. The ecology of medical care revisited. *N Engl J Med*. 2001;344(26):2021-2025. doi: 10.1056/NEJM200106283442611.
15. Ballering AV, Bonvanie IJ, Olde Hartman TC, Monden R, Rosmalen JGM. Gender and sex independently associate with common somatic symptoms and lifetime prevalence of chronic disease. *Social Science & Medicine*. 2020;253:112968. doi: 10.1016/j.socscimed.2020.112968.
16. Loikas D, Karlsson L, von Euler M, Hallgren K, Schenck-Gustafsson K, Bastholm Rahmner P. Does patient's sex influence treatment in primary care? experiences and expressed knowledge among physicians--a qualitative study. *BMC family practice*. 2015;16:137. doi: 10.1186/s12875-015-0351-5.

17. Klijs B, Scholtens S, Mandemakers JJ, Snieder H, Stolk RP, Smidt N. Representativeness of the LifeLines cohort study. *PLoS ONE*. 2015;10(9):e0137203. doi: 10.1371/journal.pone.0137203.
18. Scholtens S, Smidt N, Swertz MA, et al. Cohort profile: LifeLines, a three-generation cohort study and biobank. *International Journal of Epidemiology*. 2015;44(4):1172-80. doi: 10.1093/ije/dyu229.
19. Warmink-Perdijk WDB, Peters LL, Tigchelaar EF, et al. Lifelines NEXT: A prospective birth cohort adding the next generation to the three-generation lifelines cohort study. *European Journal of Epidemiology*. 2020;35(2):157-168. doi: 10.1007/s10654-020-00614-7.
20. McIntyre K, Lanting P, Deelen P, et al. Lifelines COVID-19 cohort: Investigating COVID-19 infection and its health and societal impacts in a dutch population-based cohort. *BMJ Open*. 2021;11:e044474. doi: 10.1136/bmjopen-2020-044474.
21. National Institute for Health and the Environment, (RIVM). Development of COVID-19 in graphs. https://www.rivm.nl/coronavirus-covid-19/grafieken. Updated 2020. Accessed 25 November, 2020.
22. Brandsma E, Verhagen HJMP, van de Laar T,J.W., Claas ECJ, Cornelissen M, van den Akker E. Rapid, sensitive, and specific severe acute respiratory syndrome coronavirus 2 detection: A multicenter comparison between standard quantitative reverse-transcriptase polymerase chain reaction and CRISPR-based DETECTR. *The Journal of infectious diseases*. 2021;223(2):206-213. doi: 0.1093/infdis/jiaa641.
23. van Zon SKR, Snieder H, Bültmann U, Reijneveld SA. The interaction of socioeconomic position and type 2 diabetes mellitus family history: A cross-sectional analysis of the lifelines cohort and biobank study. *BMJ Open*. 2017;7(4):e015275. doi: 10.1136/bmjopen-2016-015275.
24. Menni C, Valdes AM, Freidin MB, et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. *Nature Medicine*. 2020;26:1037-1040. doi: 10.1038/s41591-020-0916-2.
25. Klein SL, Dhakal S, Ursin RL, Deshpande S, Sandberg K, Mauvais-Jarvis F. Biological sex impacts COVID-19 outcomes. *PLoS Pathogens*. 2020;16(6):e1008570. doi: 10.1371/journal.ppat.1008570.
26. Peckham H, de Gruijter NM, Raine C, et al. Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ITU admission. *Nature Communications*. 2020;11(1):6317. doi: 10.1038/s41467-020-19741-6.
27. Sikkema RS, Pas SD, Nieuwenhuijse DF, et al. COVID-19 in health-care workers in three hospitals in the south of the netherlands: A cross-sectional study. *The Lancet infectious diseases*. 2020;20(11):1273-1280. doi: 10.1016/S1473-3099(20)30527-2.
28. Schwartz KL, Achonu C, Buchan SA, et al. COVID-19 infections among healthcare workers and transmission within households. *medrxiv*. 2020. doi: 10.1101/2020.06.12.20129619.
29. Eyre DW, Lumley SF, O'Donnell D, et al. Differential occupational risks to healthcare workers from SARS-CoV-2 observed during a prospective observational study. *eLife*. 2020;9. doi: 10.7554/eLife.60675.
30. Groeneveld JM, Ballering AV, van Boven K, Akkermans RP, Olde Hartman TC, Uijen AA. Sex differences in incidence of respiratory symptoms and management by general practitioners. *Family practice*. 2020;37(5):631-636. doi: 10.1093/fampra/cmaa040.
31. Moran KR, Del Valle SY. A meta-analysis of the association between gender and protective behaviors in response to respiratory epidemics and pandemics. *PloS ONE*. 2016;11(10):e0164541. doi: 10.1371/journal.pone.0164541.
32. Abate BB, Kassie AM, Kassaw MW, Aragie TG, Masresha SA. Sex difference in coronavirus disease (COVID-19): A systematic review and meta-analysis. *BMJ open*. 2020;10(10):e040129. doi: 10.1136/bmjopen-2020-040129.

33. Dehingia N, Raj A. Sex differences in COVID-19 case fatality: Do we know enough? *The Lancet Global Health*. 2021;9(1):e14-e15. doi: 10.1016/S2214-109X(20)30464-2.
34. Gerritsen DL, Oude Voshaar RC. The effects of the COVID-19 virus on mental healthcare for older people in the netherlands. *International Psychogeriatrics*. 2020;32(11):1353-1356. doi: 10.1017/S1041610220001040.

Appendices

Appendix A. Participants per measurement, stratified by sex

	Participants, N (%)		New unique participants, N (%)	
	Male	**Female**	**Male**	**Female**
Measurement 1	20,571 (38.7)	32,556 (61.3)	20,571 (38.7)	32,556 (61.3)
Measurement 2	19,591 (38.7)	31,030 (61.3)	2,049 (40.6)	2,992 (59.4)
Measurement 3	19,152 (38.7)	30,322 (61.3)	960 (41.5)	1,352 (58.5)
Measurement 4	18,295 (38.8)	28,892 (61.2)	467 (42.3)	637 (57.7)
Measurement 5	17,651 (38.9)	27,712 (61.1)	353 (42.0)	488 (58.0)
Measurement 6	16,629 (38.8)	26,178 (61.2)	182 (40.1)	272 (59.9)
Measurement 7	16,647 (38.8)	26,203 (61.2)	3,590 (39.8)	5,435 (60.2)
Measurement 8	15,060 (39.4)	23,126 (60.6)	194 (42.4)	264 (57.6)
Measurement 9	13,726 (39.0)	21,444 (61.0)	152 (39.5)	233 (60.5)
Measurement 10	12,536 (38.1)	20,408 (61.9)	97 (31.9)	207 (68.1)
Measurement 11	13,442 (38.3)	21,617 (61.7)	205 (36.6)	355 (63.4)
Measurement 12	13,892 (38.4)	22,242 (61.6)	253 (41.3)	359 (58.7)
Measurement 13	13,558 (38.6)	21,597 (61.4)	200 (40.1)	299 (59.9)
Total	210,750 (38.7)	333,327 (61.3)	29,273 (39.2)	45,449 (60.8)

Appendix B. Characteristics of participant's infected with COVID-19, stratified by method of diagnosis.

Characteristics of the population with COVID-19, split by method of diagnosis			
		Positive PCR test (N=183)	**Physician's COVID-19 diagnosis (N=983)**
Sex, N (%)	Female	131 (71.6%)	661 (67.2%)
	Male	52 (28.4%)	322 (32.8%)
Age, mean (SD)		51.5 (10.3)	51.2 (12.1)
Education, N (%)	Low	17 (9.3%)	87 (8.9%)
	Medium	100 (54.6%)	500 (50.9%)
	High	34.3 (35.0%)	386 (39.3%)
Frequent handwashing and use of desinfectant, N (%)		176 (96.2%)	959 (97.6%)
Social distancing, N (%)		173 (94.5%)	973 (99.0%)
Avoidance of public transport, N (%)		127 (69.4%)	778 (79.1%)
Covering nose and mouth in public, N (%)		31 (16.9%)	137 (13.9%)
Contact profession, N (%)	No	61 (33.3%)	476 (48.4%)
	Yes	61 (33.3%)	217 (22.1%)
	Yes, in education	11 (6.0%)	96 (9.8%)
	Yes, in healthcare	28 (15.3%)	74 (7.5%)
Smoking, N(%)		<10 (<5.0%)	86 (8.7%)

Appendix C. Adjustment for somatic diseases

Overarching disease type	Examples given in Lifelines questionnaire[a]
Cardiovascular disease	High blood pressure
	Heart attack
	Narrowing of the arteries in the legs
	Stroke/TIA
	Other heart and/or coronary diseases
Lung disease	Asthma
	COPD
	Chronic bronchitis
Liver disease	Cirrhosis
Kidney disease	Reduced kidney function
Diabetes	Diabetes mellitus Type 1
	Diabetes mellitus Type 2
Chronic muscle disease	MS
Auto-immune illness	Celiac disease
	Inflammatory bowel disorder
	Rheumatoid Arthritis
	Lupus
Cancer	Any form of cancer
Neurological disease	Dementia
	Parkinson's disease
	Alzheimer's disease
Problems with the spleen	Sickle cell anemia
	Removal of spleen
Other chronic health conditions	Open answer option

[a]These are examples given in the Lifelines questionnaires, participants with chronic diseases beyond these examples were able to mention these as well.

Appendix D. Bivariate analyses (n=74,722)

Dependent variable: COVID-19 diagnosis		
Predictor		**OR (95% CI)**
Female sex		**1.37 (1.21-1.55)**
Age		**0.98 (0.98-0.99)**
Educational attainment	Low	1.00 (ref)
	Medium	1.03 (0.91-1.16)
	High	**1.14 (1.01-1.28)**
Chronic disease present		**1.18 (1.04-1.35)**
Smoking		0.86 (0.69-1.07)
Frequent handwashing and use of desinfectant		**1.92 (1.34-2.74)**
Social distancing		**2.13 (1.37-3.32)**
Avoidance of public transport		**1.35 (1.17-1.56)**
Covering nose and mouth in public		**1.62 (1.36-1.92)**
Household members ≤ 18 years		0.98 (0.86-1.11)
Household members 19-59 years		0.88 (0.78-1.00)
Household members ≥ 60 years		**0.61 (0.53-0.70)**
Working from home		**1.36 (1.20-1.54)**
Contact profession	No	1.00 (ref)
	Yes	**1.58 (1.37-1.83)**
	Yes, in education	1.06 (0.86-1.30)
	Yes, in healthcare	**1.39 (1.13-1.72)**

Dependent variable: SARS-CoV-2 PCR test

Predictor		OR (95% CI)
Female sex		**2.04 (1.74-2.41)**
Age		**0.97 (0.96-0.97)**
Educational attainment	Low	1.00 (ref)
	Medium	**0.80 (0.69-0.92)**
	High	**1.47 (1.28-1.69)**
Chronic disease present		1.09 (0.88-1.34)
Smoking		0.88 (0.67-1.17)
Household members ≤ 18 years		**1.54 (1.29-1.86)**
Household members 19-59 years		**1.56 (1.28-1.91)**
Household members ≥ 60 years		**0.45 (0.36-0.57)**
Working from home		**0.29 (0.23-0.35)**
Contact profession	No	1.00 (ref)
	Yes	**1.84 (1.53-2.22)**
	Yes, in education	**0.53 (0.36-0.77)**
	Yes, in healthcare	**7.61 (6.35-9.11)**

CHAPTER 11A

Ballering A.V., Oertelt-Prigione S., olde Hartman T.C., Rosmalen J.G.M. (2022) Response to Rossato et al. *Journal of Women's Health, 31*(6), 896-898.

Response to Rossato et al.

We thank Rossato *et al.* for their interest in our work and applaud their efforts to assess sex differences in COVID-19-related mortality. We agree with the authors that sex- and gender-related differences with regards to COVID-19 became apparent during the pandemic. For example, although our work shows no sex difference in COVID-19 diagnoses and testing in the general population, we do demonstrate that female healthcare workers have significantly lower odds of being tested and diagnosed with COVID-19 than male healthcare workers.[1] However, as the authors rightfully note, age is an important modifier in the association between sex and health outcomes.[2]

Therefore, we build further upon the analyses conducted by Rossato *et al.* by assessing sex differences in hospitalization rates due to COVID-19 in the Dutch Lifelines COVID-19 Cohort Study.[3-5] As data collection proceeded, we included 20 consecutive measurements collected between March 2020 and March 2021, in contrast to our previous study in which we included 13 consecutive measurements up to August 2020.[1] In total, 76,422 participants of the general population completed 774,826 questionnaires about their mental and physical health during the COVID-19 pandemic. During this time, the alpha variant (B.1.1.7) was predominant in the Netherlands and by the end of March 2021 a minority of participants was fully vaccinated (5.2%).

The majority of the COVID-19-positive participants in our study was female (n= 2,149; 66.2% female). However, previous studies show that male participants are more frequently hospitalized than female participants.[6,7] Similarly, in our study, 41 women (1.9% of COVID-19-positive women) and 54 men (4.8% of COVID-19-positive men) were hospitalized due to COVID-19 (X^2=22.0; Df=1; p<0.001). The mean age of hospitalized women was 55.8 (SD=12.0) years, compared to 59.0 (SD=10.7) years in men. This difference is not statistically significant ($t_{(92)}$=1.06; p=0.29). Bivariate logistic regression analysis shows that women have lower odds than men for hospitalization due to COVID-19 (odds ratio [OR]=0.38; 95%CI=0.24-0.58). Multiple logistic regression analysis, in which we adjusted for participants' age and the presence of chronic disease, shows that sex (OR=0.44; 95%CI=0.29-0.68), age (OR=1.05; 95%CI=1.03-1.08) and the presence of chronic disease (OR=1.85; 95%CI=1.14-3.02) are all significantly associated with hospitalization due to COVID-19. Sex-by-age and sex-by-chronic disease interaction terms were not statistically significant (OR=1.01; 95%CI=0.96-1.06 and OR=0.77; 95%CI=0.29-2.05, respectively). This indicates that the association between age and hospitalization, as well as the association between the presence of a chronic disease and hospitalization due to COVID-19 do not differ between women and men.

The lower odds of female hospitalization due to COVID-19 are most likely attributable to both sex and gender-related factors. Women's innate and humoral immune responses, as well as their ability to balance inflammation and tissue damage, appear to be stronger than men's.[6,8] These immunological sex differences may result in more effective clearing of infection in women and a potentially accelerated recovery after infection.[9] Additionally, higher age-adjusted rates of pre-existing somatic comorbidities are reported in male COVID-19 patients, including cardiovascular disease, which associate with a poor COVID-19 prognosis.[10-12] In addition to sex-specific factors, gender-related factors may also associate with the course of SARS-CoV-2 infection. For example, men are generally more likely to display poor health-related behaviors that worsen a COVID-19 prognosis, such as smoking and poor diet.[13,14] The above-mentioned sex-related factors may interact with gender-specific factors and thus synergistically influence the prognosis of COVID-19. Notably, sex- and gender-based differences regarding hospitalization due to COVID-19 may differ between cultures and location. In general, our findings regarding a female preponderance in COVID-19 diagnoses and poor prognosis of COVID-19 are in line with those of Rossato *et al*. However, further research could focus on sex- and gender-based differences in physical consequences (e.g., post COVID-19 condition) and sociocultural consequences (e.g., loss of productivity) due to COVID-19 and the pandemic in general.

References

1. Ballering AV, Oertelt-Prigione S, LCRI, Olde Hartman TC, Rosmalen JGM. Sex and gender-related differences in COVID-19 diagnoses and SARS-CoV-2 testing practices during the first wave of the pandemic: The dutch lifelines COVID-19 cohort study. *Journal of Women's Health (Larchmt.)*. 2021;30(12):1686-1692. doi: 10.1089/jwh.2021.0226.
2. Green MS, Nitzan D, Schwartz N, Niv Y, Peer V. Sex differences in the case-fatality rates for COVID-19—A comparison of the age-related differences and consistency over seven countries. *PLoS ONE*. 2021;16(4):e0250523. doi: 10.1371/journal.pone.0250523.
3. McIntyre K, Lanting P, Deelen P, et al. Lifelines COVID-19 cohort: Investigating COVID-19 infection and its health and societal impacts in a dutch population-based cohort. *BMJ Open*. 2021;11:e044474. doi: 10.1136/bmjopen-2020-044474.
4. Scholtens S, Smidt N, Swertz MA, et al. Cohort profile: LifeLines, a three-generation cohort study and biobank. *International Journal of Epidemiology*. 2015;44(4):1172-80. doi: 10.1093/ije/dyu229.
5. Klijs B, Scholtens S, Mandemakers JJ, Snieder H, Stolk RP, Smidt N. Representativeness of the LifeLines cohort study. *PLoS ONE*. 2015;10(9):e0137203. doi: 10.1371/journal.pone.0137203.
6. Nachtigall I, Bonsignore M, Thürmann P, et al. Sex differences in clinical course and intensive care unit admission in a national cohort of hospitalized patients with COVID-19. *Journal of Clinical Medicine*. 2021;10(21):4954. doi: 10.3390/jcm10214954.
7. Peckham H, de Gruijter NM, Raine C, et al. Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ITU admission. *Nature Communications*. 2020;11(1):6317. doi: 10.1038/s41467-020-19741-6.
8. Scully EP, Haverfield J, Ursin RL, Tannenbaum C, Klein SL. Considering how biological sex impacts immune responses and COVID-19 outcomes. *Nature Reviews Immunology*. 2020;20(7):442-447. doi: 10.1038/s41577-020-0348-8.
9. Bischof E, Oertelt-Prigione S, Morgan R, Klein SL, The Sex and Gender in COVID19 Clinical Trials Working Group (SGC), Gender and COVID19 Working Group. Towards precision medicine: Inclusion of sex and gender aspects in COVID-19 clinical studies-acting now before it is too late-A joint call for action. *International Journal of Environmental Research and Public Health*. 2020;17(10):3715. doi: 10.3390/ijerph17103715.
10. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in wuhan, china: A retrospective cohort study. *The Lancet*. 2020;395(10229):1054-1062. doi: 10.1016/S0140-6736(20)30566-3.
11. Grasselli G, Zangrillo A, Zanella A, et al. Baseline characteristics and outcomes of 1591 patients infected with SARS-CoV-2 admitted to ICUs of the lombardy region, italy. *JAMA*. 2020;323(16):1574-1581.
12. Petrilli CM, Jones SA, Yang J, et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in new york city: Prospective cohort study. *BMJ*. 2020;369:m1966. doi: 10.1136/bmj.m1966.
13. Haitao T, Vermunt JV, Abeykoon J, et al. COVID-19 and sex differences: Mechanisms and biomarkers. *Mayo Clinic Proceedings*. 2020;95(10):2189-2203. doi: 10.1016/j.mayocp.2020.07.024.
14. Ahmed SB, Dumanski SM. Sex, gender and COVID-19: A call to action. *Canadian journal of public health*. 2020;111(6):980-983. doi: 10.17269/s41997-020-00417-z.

PART 5

Adequate conceptual assessment
and implementation of knowledge

CHAPTER 12

Ballering, A.V., Burke, S.M., Maeckelberghe, E.L.M., & Rosmalen, J.G.M. (2023). How to Ensure Inclusivity in Large-Scale General Population Cohort Studies? Lessons Learned with Regard to Including and Assessing Sex, Gender, and Sexual Orientation. *Archives of Sexual Behavior, 52,* 2163-2172

Abstract

Despite recent advances in the measurement of sex, gender and, sexual orientation in large-scale cohort studies, the three concepts are still gaining relatively little attention, may be mistakenly equated, or non-informatively operationalized. The resulting imprecise or lacking information hereon in studies is problematic, as sex, gender, and sexual orientation are important health-related factors. Omission of these concepts from general population cohort studies might dismiss participants' identity and experiences, and pushes research on sexual or gender minority populations towards purposive sampling, potentially introducing selection bias. It also reinforces the unintentional notion of irrelevance of these concepts to health research, ultimately disadvantaging sexual and gender minority populations. Similarly, a lack of uniform measures on sex, gender, and sexual orientation hampers multi-cohort studies in which data from multiple studies are combined, facilitating increased statistical power. This paper discusses the encountered pitfalls and lessons learned on including and assessing sex, gender, and sexual orientation in large-scale general population cohort studies, exemplified by the Dutch Lifelines Cohort Study. Additionally, we propose hands-on strategies on how to operationalize these concepts in an inclusive manner that is useful for large-scale general population cohort studies.

Introduction

Although many advocates have been arguing for the inclusion of sex, gender, and sexual orientation in health research for decades, it has only been since the late 2000s that this movement gained momentum in epidemiological cohort studies.[1] In the slipstream of the increased attention for these concepts, we initiated an epidemiological research project to assess the associations between sex, gender, and common somatic symptoms in Lifelines. Lifelines is a large general population cohort study, with a three-generation design including over 167,000 participants from the North of the Netherlands. However, when embarking upon this project, we realized that our intended dataset did not include sufficient information to adequately answer our research questions. No information on participants' gender, sex assigned at birth, or sexual orientation had been included during the data collection.

This lack of precise and valid information on sex, gender, and sexual orientation is not a stand-alone occurrence, but similar to other general population cohort studies.[2] Two leading large-scale cohort studies, The UK Biobank and HUNT[a], do not register any dimension of gender, while a third large-scale registry, the Veterans Health Administration (VHA), does register self-reported, categorized gender identity, albeit not routinely for all participants.[3-6] These three studies or registries all derive sex from central registries, such as birth certificates. The UK biobank complements their sex variable with genetic sex, and does allow for adaptations in participants' sex, rendering the sex variable a mix between recorded and self-reported sex. Similarly, the VHA allows for adjustment of birth sex. Sexual orientation is differently assessed in these studies, with the UK Biobank assessing lifetime number of same-sex sexual partners, while HUNT and VHA assesses self-reported sexual identity.

These examples align with the recent evaluation of the National Academies of Science, Engineering, and Medicine of sex, gender identity, and sexual orientation measures in research, administrative, and clinical contexts:

"This evaluation revealed not only how much progress has been made in the development and refinement of sex, gender identity and sexual orientation measures that identify sexual and gender minority populations, but also how much progress remains to be made. Although measures [...] become more widely implemented in data collection efforts, few of the measures in use are explicitly inclusive of gender identities that lie outside of the gender binary and many continue to rely on terminology or language that is considered invalidating or offensive to some sexual and gender minorities." (p. IX)[7]

Thus, although increasing attention has been directed towards including and assessing sex, gender, and sexual orientation inclusively over the past decades, many leading large-scale cohort studies still use insufficient measures for these concepts. Furthermore, the lack of uniform measures on sex, gender (identity), and sexual orientation hampers multi-cohort studies on these concepts, in which data derived from multiple cohort studies can be combined to facilitate increased statistical power.

The paucity of information about participants' sex, gender, and sexual orientation in general population cohorts is problematic, as over time a growing body of evidence has shown that these variables are important factors in health and disease.[8,9] Some health problems, for example, occur more frequently in women than in men, either largely due to their biological sex (e.g., breast cancer), or due to an interaction between sex and gender, for example in osteoarthrosis in which both hormonal levels and occupational hazards play a role.[10] Additionally, literature shows that the transgender and gender diverse (TGD) population, or people with a lesbian, gay. or bisexual (LGB) sexual orientation are more at risk for chronic somatic diseases and psychiatric disorders (e.g., because of minority stress and related (mental) health problems).[11] Omission of sex, gender, and sexual orientation in studies also reinforces the unintentional notion of irrelevance of these concepts to health research.

To fully grasp the necessity to include sex, gender, and sexual orientation in health research, it is important to clarify the differences between the concepts. For example, both sex and gender have been regarded for a long time as dichotomous, synonymic concepts that can function as a proxy measure for each other, despite the two being different concepts.[7] Similarly, sexual orientation is a concept that is distinctly different from sex and gender.[12] Therefore, we provide extensive definitions of the three concepts and their concomitant dimensions in ***Table 1***.

Although a fundamental variable such as participants' sex is usually included in cohort studies, albeit sometimes inaccurately assessed, variables on any dimension of gender, including gender identity, and sexual orientation may be omitted by design in cohort studies. Possibly, researchers are unaware of these topics and the concomitant multidimensionality, or the omission could stem from the researchers' idea of these being supposedly sensitive questions with which participants should not be confronted, as this could potentially result in reduced retention.[19] Furthermore, knowledge on how to assess in a sensitive, yet informative manner participants' biological sex, gender, and sexual orientation is lacking.[9]

Table 1. Definitions and Dimensions of Sex, Gender, and Sexual Orientation

Sex assigned at birth and intersex variations	A biological construct that encompasses the biology, among others genes, hormones, physiology and anatomy, of female and male bodies. Sex is usually assigned at birth. Although sex is often seen as a female/male binary, sex characteristics exist on a continuum, thus challenging the dichotomous beliefs about biological sex.[13] Intersex variations include a wide range of innate differences that relate to gonads, chromosomes, and genitals that do not fit the typical medical or social binary norms for female and male bodies.[14] Strict definitions of intersex variations include chromosomal variations of the sex-chromosomes (e.g., Klinefelter syndrome), genetic mutation(s) resulting in hormonal disturbances that affect sexual development (e.g., androgen insensitivity syndrome), or variations of the internal and/or external genital organs (e.g., Mayer-Rokitansky-Küster-Hauser syndrome).[15] A more liberal definition of intersex variations also includes common variations of external genital organs, such as hypospadias or cryptorchism for which an operation was required.
Gender	The embodiment of different roles, behaviors, identities, and relationships of men and women prescribed by societal norms, which results in different expectations, opportunities, and experiences of men and women in a given society.[13] Gender comprises multiple dimensions[13]: 1. Gender identity: Describes whether people identify themselves as women, men, non-binary or another gender. Gender identity in itself is multidimensional, including[16]: a) Felt-gender: the extent to which one experiences (in)congruence between the feeling of being a woman or a man, and one's sex assigned at birth. b) Gender contentedness: the degree of satisfaction or dysphoria one experiences with regards to their gender. c) Conformity of gender expression: the degree of compliance with gender-related norms, such as expressing gender via hobbies or clothing. 2. Gender roles: The behavioral norms applied to people based on societal expectations and mores related to their gender. 3. Gender relations: Encompasses the interactions between people based on their ascribed gender. 4. Institutionalized gender: Refers to how power is distributed between genders in institutions.
Sexual orientation	The American Psychological Association[17] refers to sexual orientation as the description of people's enduring pattern of emotional, romantic and/or sexual attractions and preferences based on their sex and gender relative to the sex and gender of (a) potential partner(s), and people's sense of identity based on those attractions, related behaviors and membership in a community of others who share those attractions. Sexual orientation is multidimensional and can be approached from a sex perspective, a gender perspective and a combined sex/gender perspective.[11,18] The dimensions of sexual orientation include: 1. Sexual identity: Refers to how people label themselves in relation to their partner's or partners' sex and/or gender preference. 2. Sexual behavior: Describes people's sexual partnered behaviors and activities. 3. Sexual attraction: Describes the sexual interests, approaches, attractions and fantasies revolving around the sex and/or gender of the chosen or desired partner(s).

Therefore, this paper aims to describe and discuss lessons learned regarding the inclusion and assessment of sex, gender, and sexual orientation in general population cohort studies. We will illustrate our points on inclusivity by using our own experiences with the assessment of sex, gender, and sexual orientation in Lifelines to show how these may be handled within general population cohort studies.[20,21] We will also propose concrete strategies to assess these concepts in cohort studies, while acknowledging that researchers are often constrained in what they can ask from participants by practicalities (i.e., costs, space, and participant burden) as well as participants' potential concerns regarding their privacy and disclosure of sex, gender, and sexual orientation. Despite their separate discussion in the text, sex, gender, and sexual orientation are intrinsically linked and their interactional effect on health will be discussed as well. Ultimately, the to-be-discussed lessons refer to the larger, overarching concept of inclusivity in large-scale data studies. However, we are aware that cultural and social mores do not always allow for a setting in which sex, gender, and sexual orientation can be openly disclosed, researched, and discussed. Therefore, the lessons described here should be interpreted with cultural and social frames of reference in mind.

Sex

Participants' sex appears to be a straightforward concept at first glance. However, in the context of health research, it is more complicated than what may be initially expected. In Lifelines, for example, participants' sex assigned at birth was derived from the municipal registry.[22] This resulted in an inconsistent operationalization of sex in two ways. First, the information provided was restricted to a female/male binary, which disregards the possibility of intersex variations. Second, for the vast majority of participants, municipally-registered sex comprises sex assigned at birth. However, as of 2014, the Dutch law allows for individuals to change their sex in the municipal registry in a more accessible manner than before.[b]

Recently, also an "X" to indicate non-binary sex was introduced. Thus, for a minority of participants who changed their sex in the municipal registry (e.g., due to strong gender incongruent feelings) municipally-registered sex may reflect their gender identity rather than their sex assigned at birth. However, although often conflated, conceptually, gender identity differs substantially from sex assigned at birth and should not be reduced to mere sex traits.

Intersex variations

Intersex variations include a wide range of innate differences that relate to gonads, chromosomes, and genitals that do not fit the typical medical or social binary norms for female and male bodies.[14] The prevalence rate of intersex variations range from 0.05% to 1.7% in the general population.[15,23] The variation in prevalence rates is reinforced as general population studies do not routinely include items that assess the presence of intersex variations and the exact definition of intersex variations remains a matter of debate.[14] Additionally, not all intersex variations are readily identified at birth, but rather later in life. However, to facilitate research exploring sex- and thus intersex-related health factors, identification of participants with an intersex variation is required.

As no specific question assessing intersex variation was included in Lifelines, complementary approaches to identify participants with an intersex variation have been previously used[c]: text fields of items assessing disorders, birth defects, and operations were searched for expressions of potential intersex variations, intersex birth variations, and gonad-related operations.[22] Upon applying a strict definition of intersex variations, a point prevalence for intersex variations of 0.05% in Lifelines was estimated, whereas a more liberal definition in which common variations of external organs such as hypospadias were included, yielded a point prevalence of 0.55% (***Table 1***). Ideally, this type of strategy should function merely as a complementary approach in addition to a specific intersex-identifying item in a survey.

Intersex variations have different etiologies. Some intersex variations have a sex-chromosomal-related etiology that can be detected by genetic approaches. In Lifelines, first-stage quality control procedures excluded participants' genetic material that did not correspond with the municipally-registered sex, as these were considered clerical or handling errors. This ultimately reduced the diversity of released data and resulted in a loss of information about intersex variations in Lifelines. This, as well as the relatively late or missed diagnosis of some intersex variations in general, likely caused Lifelines' point prevalence to be an underestimation of the true prevalence. Currently, Lifelines also identifies relatives of participants in whom a genetic and municipally-registered sex disconcordance occurs, and by using pedigree information and information provided by the family members about their relatives' sex could confirm a sample mix-up. However, many cohort studies have no multiple-generation design, and cannot assess pedigree information and familial relationships. Other large-scale cohort studies with a similar quality control pipeline, such as the UK Biobank project, did not exclude data derived from participants in whom genetically-inferred sex based on sex chromosomes

differed from self-reported sex.[4] Rather, data derived from participants with a potential intersex variation or TGD identity was indicated as such, maintaining an inclusive and diverse study population. Some intersex variations are not readily detectable by genetic screening of the sex chromosomes as described above, and in some general population cohort studies no genetic approaches are included in the design. Therefore, expanding the male/female binary of participants' sex with a non-binary option in the assessment is pivotal in obtaining more detailed data about people with intersex variations, allowing for more tailored research in this population. ***Table 2*** describes a set of survey items that allow for identification of participants with an intersex variation. People with an intersex variation may be assigned a sex at birth that reflects their sex characteristics at time of birth, which are not necessarily indicative of an intersex variation. Thus, by including "intersex" as an option when assessing participants' sex assigned at birth, inconsistent results may be obtained. Therefore, an additional item that describes intersex variations allows for the identification of intersex people in a general population cohort[7]. The item in ***Table 2*** is congruent with the current Dutch context, as of recently Dutch legislation eased the process of assigning an "X" on a birth certificate, indicating that the sex assigned at birth could not be irrefutably determined.

Table 2. Survey Items including a Non-Binary/Intersex Option

Could you indicate your sex assigned at birth, as stated on your birth certificate?
() Male (M)
() Female (F)
() Non-binary (X)
() Other, please write down your preferred term…
Were you born with a variation in sex characteristics (this is sometimes called intersex or an intersex variation)?
() Yes
() No

Gender

Many cohort studies, including Lifelines, do not include specific questions assessing any dimension of participants' gender. However, for our studies we were interested in the independent associations between gender roles and sex, and common somatic symptoms. Therefore, we recently showed how a data-driven method can be used to calculate a composite gender index based on participants' gendered psychosocial characteristics for cohorts that lack data on gender.[22] We defined a gender score that quantified participants' adherence to feminine and masculine psychosocial

characteristics including but not limited to hobbies, personality traits, type of profession, time spend on household activities, and dietary preferences. As a result, participants were placed on a continuum ranging from 0%, i.e., fully masculine, to 100%, i.e., fully feminine.

The method is suitable for general population cohort studies that lack measures on gender, and facilitates a gender measure specific to the context of the study. A strong advantage of this measure is that it is sensitive to the time, place, and society-bound nature of gender roles. Other existing measures, including the Bem Sex Role Inventory (BSRI), have been criticized and are argued to hold limited validity to operationalize femininity and masculinity.[24-26] These instruments measure gender via items that stereotype masculine and feminine traits, while gender roles are a broad concept that is largely dependent on the respective time, place, and society.[22]

Gender measures based on previously collected survey data usually assess gender roles and/or gender relations, and cannot capture participants' current gender identity (***Table 1***).[22,27,28] Although gender roles and gender identity are intrinsically linked, a gender identity measure, in contrast to a gender role measure, cannot be calculated after data collection. Yet, ideally, both participants' gender roles and gender identity are assessed in cohort studies as both dimensions are known to affect TGD and cisgender participants' health substantially.[13,22,29]

Gender identity is a fluid, continuous, and multidimensional concept. The embodiment and expression of gender identity may differentiate over time, especially in adolescents,[2,30] allowing for fluidity of gender identity to be captured by a repeated measures design. The continuous nature of the concept can be captured by assessing participants' feminine or masculine identity on unipolar two-dimensional continuous scales. This allows for measuring the extent of participants' adherence to gender identities. Preferably, gender identity should be assessed via at least a two-step approach, in which assessment of one's sex assigned at birth and current gender identity are combined (***Table 3***). This allows for identification of participants with gender incongruent feelings.[d]

Gender identity is multidimensional and thus multiple, interlinking domains together define one's gender identity. Building further on the initial model for the multidimensionality of gender identity,[31] studies refined the dimensions.[16,32] Recent studies, for example, include (1) felt-gender, (2) gender contentedness, and (3) gender conformity (***Table 1***).[32,33] The multidimensionality of gender identity calls for an approach that moves beyond the common two-step approach that merely combines sex assigned

at birth and current gender identity. Previous studies, that assessed a dimensional approach to gender identity and gender incongruency, have proven the validity of multi-item questionnaires in both adults and adolescents.[34,35] However, as the number of items that can be included in general population cohort studies is limited due to the supposed burden for participants, as well as space and cost considerations, including items on the multiple dimensions of gender identity may be more feasible in smaller add-on studies that have a specific focus on gender identity in relation to health.

Table 3. Items Included in Lifelines, Based on the Two-Item Approach Combining Sex Assigned at Birth and Current Gender Identity, with Quantification of Missing Data

	Municipally-registered female participants (N=31,058)	**Municipally-registered male participants (N=21,588)**
	Missing, N (%)	**Missing, N (%)**
Most people are born as either a man or a woman and they feel comfortable in a male or female body, respectively. However, this is not the case for everyone. Some people consider themselves a man, but were born in a female body or vice versa. Some people consider themselves neither a man nor a woman. Could you indicate which statement fits your experience best?[a]		
Could you indicate which statement fits your experience best? () My sex assigned at birth was female and I currently consider myself a woman. () My sex assigned at birth was male and I currently consider myself a man. () My sex assigned at birth was female and I currently consider myself a man. () My sex assigned at birth was male and I currently consider myself a woman. () Different, namely …	104 (0.3%)	56 (0.3%)
Some people, both men and women, consider themselves masculine, for example because they have characteristics or hobbies that most people consider masculine. On the other hand, both men and women may consider themselves feminine, because they have characteristics or hobbies that most consider feminine. Some people consider themselves neither masculine nor feminine. Could you indicate, on a scale ranging from 1 to 10 in which 1 equals *strongly disagree* and 10 equals *strongly agree*, to what extent you consider yourself feminine or masculine? Please complete both questions.		
I consider myself feminine. (1) Strongly disagree ‹‹‹‹‹‹‹ ››››››› (10) Strongly agree	110 (0.4%)	272 (1.3%)
I consider myself masculine. (1) Strongly disagree ‹‹‹‹‹‹‹ ››››››› (10) Strongly agree	338 (1.1%)	95 (0.4%)

[a]Potentially, the item could be expanded with additional answer options in which a non-binary gender identity is included.

Despite the potential stigma that may surround non-cisgender identities, it has been shown that the common two-step approach is easily understandable, well accepted, and causes little to no resistance in both cisgender and TGD participants in cohort studies.[36,37] Lifelines recently included an assessment of gender adapted from the two-step approach. This assessment includes gender identity and gender roles and was reviewed by a participant panel, including TGD participants, before implementation. Importantly, out of 52,646 adult Lifelines participants, only 0.3%-1.3% of the male participants and 0.3%-1.1% of female participants did not answer these questions (***Table 3***). This indicates that the vast majority of participants was willing to complete this item.

Sexual Orientation

Akin to gender identity, specific questions regarding sexual orientation are frequently omitted in general population cohort studies. In Lifelines, for example, merely the binary sex of participants' current partner is assessed. This is only an indirect measure from which participants' sexual orientation could be inferred. If information on participants' sex assigned at birth and current gender identity is unknown, the information obtained by this item is even more multi-interpretable.

Notably, the way in which questions on sexual orientation are phrased could influence the distribution of sexual orientation in a study sample.[38] There is no generalizable rule about how items on sexual orientation should be phrased. Partly, this relates to the ongoing debate on the central axis around which sexual orientation revolves. Does one's sexual orientation revolve around the partner's sex, gender, or both? As Van Anders states:

"For example, if one is sexually attracted to men, is one attracted to penises? Social identities? Body frames? Interactions? And, how is sexual orientation defined if one is attracted to masculinity regardless of the sex of the person presenting or embodying it?" (p. 1177)[18]

Some theorize, however, that sexual orientation relates to additional concepts beyond potential partner's sex and/or gender, such as partner number and partner age.[18]

Sexual orientation is also a multidimensional concept, with three separate dimensions: sexual identity, sexual behavior and sexual attraction[e] (***Table 1***).[11,18] The apparent relevance of asking for participants' sexual orientation correlates directly to participants'

willingness to complete items on sexual orientation in relation to health.[39] Yet, not all dimensions of sexual orientation are relevant to assess in every setting. Whether or not it is appropriate and relevant to ask participants about a dimension of sexual orientation depends on the context and research question. For example, during a consult with their GP, patients may be more aware of their sexual behavior influencing their health, and they more readily disclose such information.[39] In this case, information about sexual orientation is of direct importance to people's own health. Similarly, when donating blood, it is clearly explained why the survey administered during the intake asks for donor's sexual behavior. information about one's sexual behavior may be of direct importance to transfusion safety. These examples illustrate people's willingness to disclose information on sexual behavior, as long as the rationale for assessing it is clear to participants. In large-scale cohort studies, it is necessary as well to clearly explain the health-related relevance underlying items on sexual orientation and to explain that sexual orientation may associate with the development of both psychological and physical health conditions,[40,41] and that knowledge hereon is important for public health.

Many general population cohort studies assess sexual orientation by merely asking about participants' sexual identity in terms of lesbian/gay, straight (i.e., not gay or lesbian) or bisexual, while it has been recently recommended to move beyond mere self-reported identity and to include sexual attraction and possibly behavior as well.[7] First, although self-reported identity measures allow for a relatively easy-to-analyze outcome measure, it may enforce oversimplified categorization of participants' sexual orientation. Second, it cannot explicate the central axis of a participant's sexual orientation and an asexual option is frequently overlooked. Third, such self-reported sexual identity items may cause confusion for TGD participants, as they may not know whether to reason from their sex assigned at birth or current gender identity. Even among researchers no consensus exists on whether sex assigned at birth or current gender identity should be used as reference to define sexual orientation,[42,43] rendering sexual identity items multi-interpretable. Fourth, sexual identity (and behavior) may be strongly constrained by local mores and culture and may not fully reflect participants' sexual orientation (e.g., in conservative religious communities). Last, sexual attraction underlies and complements behavior and identity, rather than behavior and identity underlying sexual attraction.[44]

To at least partly overcome these disadvantages of solely assessing sexual identity, we do not argue to abandon a self-reported sexual identity item in large-scale population cohort studies. We rather argue for complementing such an identity item with gynephilia and androphilia items. This provides an option for assessing sexual orientation that facilitates flexibility, yet restricts the answer options in such a way

that they remain meaningful for general population cohort studies such as Lifelines (***Table 4***), in which currently no items assess sexual orientation adequately. We propose to complement a self-reported sexual identity item with two unidimensional scales on which participants can indicate their sexual attraction in terms of gynephilia or androphilia in general population cohort studies. By using two unipolar scales, participants' sexual orientation can be approached and analyzed in a continuous manner, disregarding the need for categorization. It allows for participants to indicate asexuality, degree of same-sex sexual attraction, and degree of other-sex attraction.

Table 4. Proposed Survey Items on Sexual Orientation for General Population Cohort Studies.

Could you indicate on the scales below what describes you best?
I am sexually attracted to men.
(1) not at all <<<<<<<<<< >>>>>>>>>> (10) Very strongly
I am sexually attracted to women.
(1) not at all <<<<<<<<<< >>>>>>>>>> (10) Very strongly

For specific research questions related to sexual orientation and add-on studies, this item could potentially be extended by two additional unipolar scales on which participants can indicate their sexual attraction with a gender focus, instead of a sex focus. However, as the distinction between sex and gender frequently remains unclear among the general population, we propose to only assess sexual orientation with a focus on sex in general population cohorts as items with a gender focus may be misinterpreted by participants. To avoid confusion about sexual orientation in TGD or people with an intersex variation, items on sexual orientation should be combined with questions about people's sex assigned at birth and current gender identity (***Table 2*** and ***Table 3***). To the best of our knowledge, assessing sexual orientation on two unipolar scales referring to gynephilia and androphilia has not been implemented in large-scale general population cohort studies. Future research is needed to validate these items and to compare their results with those obtained by items assessing self-reported, categorized sexual identity.

Towards an Inclusive Future of Research

Here, we have described and discussed our experiences and lessons learned regarding sex, gender, and sexual orientation in large-scale general population cohort studies. We have described pitfalls in assessing and including these concepts and we have proposed strategies to operationalize these in an inclusive manner relevant for the research question at hand.

It should be emphasized that obtaining detailed information about participants' sex, gender, and sexual orientation in general population cohort studies is pivotal. First, disregarding these variables in general population cohort studies excludes the possibility of conducting studies within TGD and LGB subpopulations in a general population cohort, especially since the large study populations of cohort studies potentially allow for identification of a relatively large TGD and LGB subpopulation herein as well. As a result, studies focusing on TGD and LGB populations are usually pushed towards convenience and purposive sampling, potentially introducing selection bias.[38] Ultimately, this results in a decreased external validity of study results.[45] Nevertheless, selection bias cannot be fully dismissed in general population cohort studies either, as TGD and LGB populations may conceal aspects of their sex, gender, and/or sexual orientation,[46] resulting in an underrepresentation of TGD and LGB populations and non-random misclassification of sexual and gender minority populations potentially decreasing the validity of research findings. Second, excluding detailed information on sex, gender, and sexual orientation from general population cohort studies reinforces the current status quo in which sexual and gender minority populations are disadvantaged. Third, health-related research focusing specifically on TGD and LGB populations may ultimately contribute to better healthcare and health outcomes for these populations (e.g., by designing more personalized health interventions). Especially large general population cohort studies have the potential to identify new or more complex associations between risk factors and health of sexual and gender minority populations, but this requires adequate identification of participants' sex, gender, and sexual orientation.

Nevertheless, we acknowledge that researchers are constrained by practicalities (e.g., costs and participant burden) in what they can ask from participants: the number of items in a survey, their contents, and wording should be carefully balanced. Therefore, it follows that questions about sex, gender, and sexual orientation should be tailored to the specific setting and goal of the research.[11] Also, survey items that deviate from the sex/gender binary or heteronormative stance may cause resistance in relatively few participants.[33,47] On the other hand, omitting survey items that deviate from these norms may feel like a denial of participants' identity or lived experiences to those who identify beyond these norms.[48,49]

Furthermore, purposeful omission of survey items that go beyond dichotomous sex, gender identity, and sexual orientation, or purposeful inclusion of dichotomous items, is a normative assumption in itself: researchers should not automatically assume that participants refuse to answer these items. In contrast, recent evidence shows that participants often appreciate being able to share information about these topics[19,50,51] provided that it is clear to participants that their information is handled in compliance

with local institutional and legal privacy guidelines aimed at, among others, avoiding re-identification of anonymized or pseudonymized participants. We strongly feel that the rather small chance of resistance does not outweigh omission of inclusive items, if these allow researchers to assess and possibly aid in improving the health and empowerment of disadvantaged sexual and gender minority populations. However, to ensure acceptance as much as possible, survey items and explanatory notes on sex, gender, and sexual orientation should be implemented in collaboration with a diverse participant panel. Similarly, the collaboration with a diverse participant panel may allow for a reduction in participants' potential concealment of sex, gender, and sexual orientation.

In conclusion, to ensure inclusivity in large-scale general population cohort studies, researchers and participants, need to understand the relevance, but also the nuances and multidimensionality of participants' sex, gender, and sexual orientation. Accounting for the lessons learned described here is a step towards an inclusive future of research, but to achieve optimal inclusivity, awareness about these concepts and their interconnectedness should be routinely ingrained in the design of general population cohort studies.

Endnotes

[a] HUNT is an acronym for the Norwegian name of the study "HelseUndersøkelsen i Nord-Trøndelag".

[b] Although it is possible to legally change your sex since 1985 in a Dutch municipal registry, several criteria, including prove of a gender confirming surgery, were necessary. Therefore, TGD that did not undergo surgery could not change their sex in the official registries before 2014.

[c] A similar strategy as described here was used to identify TGD participants in Lifelines, as no specific question on gender identity and gender contentedness was included in Lifelines until 2020.

[d] Researchers may consider to add an "() I prefer not to disclose option". We decided here to exclude this option, as in general population cohort studies supposedly sensitive information, for example on traumatic experiences, is assessed as well, without providing an "I prefer not to disclose" option. If researchers wish to add this answer option, it should be added consistently throughout all items in the survey to avoid a reinforcement of the notion that information about sex, gender, and sexual orientation is more sensitive than other information.

[e] Sexual orientation is sometimes regarded as a subdimension in itself [18], equated with what we call sexual attraction in this text. For reasons of clarity we use the term sexual attraction, instead of sexual orientation, to describe the sexual interests, approaches and fantasies revolving around the sex and/or gender of one's chosen partner(s).

All Lifelines participants have provided written consent. Lifelines is performed according to the principles of the Declaration of Helsinki and is approved by the Medical Ethical Committee of the University Medical Center Groningen (number: 2007/152).

References

1. Klinge I. Gender perspectives in european research. *Pharmacological research*. 2008;58(3):183-189. doi: 10.1016/j.phrs.2008.07.011.
2. Westbrook L, Saperstein A. New categories are not enough: Rethinking the measurement of sex and gender in social surveys. *Gender & society*. 2015;29(4):534-560. doi: 10.1177/0891243215584758.
3. Åsvold BO, Langhammer A, Rehn TA, et al. Cohort profile update: The HUNT study, norway. *International Journal of Epidemiology*. 2022;52(1):e80-e91. doi: 10.1093/ije/dyac095.
4. Bycroft C, Freeman C, Petkova D, et al. The UK biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562(7726):203-209. doi: 10.1038/s41586-018-0579-z.
5. Grozdanic T, Lamba S, Moy E, et al. Primary healthcare experiences of lesbian, gay, and bisexual veterans. *Office of Health Equity, Veterans Health Administration, U.S. Department of Veterans Affairs*. 2022.
6. Veterans Health Administration. Transgender and gender diverse veteran healthcare: Birth sex and gender identity. https://www.patientcare.va.gov/LGBT/docs/2022/Birth-Sex-Gender-Identity-FactSheet-for-Veterans-2022.pdf. Updated 2022. Accessed 11 November, 2022.
7. National Academies of Sciences, Engineering, and Medicine. Measuring sex, gender identity, and sexual orientation. *National Acadmies of Sciences, Engineering, and Medicine*. 2022;26424.
8. Regitz-Zagrosek V. Sex and gender differences in health. *EMBO reports*. 2012;13(7):596-603. doi: 10.1038/embor.2012.87.
9. Bränström R, van der Star A, Pachankis J. European-wide monitoring of sexual and gender minority health: A necessary and feasible task for advancing health equity. *European journal of public health*. 2019;29(2):199-200. doi: 10.1093/eurpub/ckz007.
10. Laitner MH, Erickson LC, Ortman E. Understanding the impact of sex and gender in osteoarthritis: Assessing research gaps and unmet needs. *Journal of women's health*. 2021;30(5):634-641. doi: 10.1089/jwh.2020.8828.
11. Salomaa AC, Matsick JL. Carving sexuality at its joints: Defining sexual orientation in research and clinical practice. *Psychological assessment*. 2019;31(2):167-180. doi: 10.1037/pas0000656.
12. Fausto-Sterling A. Gender/sex, sexual orientation, and identity are in the body: How did they get there? *Journal of sex research*. 2019;56(4-5):529-555. doi: 10.1080/00224499.2019.1581883.
13. Johnson JL, Greaves L, Repta R. Better science with sex and gender: Facilitating the use of a sex and gender-based analysis in health research. *International Journal for Equity in Health*. 2009;8(1):14. doi: 10.1186/1475-9276-8-14.
14. Rosenwohl-Mack A, Tamar-Mattis S, Baratz AB, et al. A national study on the physical and mental health of intersex adults in the U.S. *PLoS ONE*. 2020;15(10):e0240088. doi: 10.1371/journal.pone.0240088.
15. Blackless M, Charuvastra A, Derryck A, Fausto-Sterling A, Lauzanne K, Lee E. How sexually dimorphic are we? review and synthesis. *American Journal of Human Biology*. 2000;12(2):151-166.
16. Burke SM. Commentary: Gender diversity and adolescent mental health- A reflection on potter et al. (2020). *Journal of child psychology and psychiatry*. 2021;62(2):180-183. doi: 10.1111/jcpp.13309.
17. American Psychological Association. Answers to your questions for a better understanding of sexual orientation and homosexuality. http://www.apa.org/topics/lgbtq/orientation. Updated 2008. Accessed 10 November, 2022.

18. van Anders SM. Beyond sexual orientation: Integrating gender/sex and diverse sexualities via sexual configurations theory. *Archives of sexual behavior*. 2015;44(5):1177-1213. doi: 10.1007/s10508-015-0490-8.
19. Sell RL. Challenges and solutions to collecting sexual orientation and gender identity data. *American Journal of Public Health*. 2017;107(8):1212-1214. doi: 10.2105/AJPH.2017.303917.
20. Klijs B, Scholtens S, Mandemakers JJ, Snieder H, Stolk RP, Smidt N. Representativeness of the LifeLines cohort study. *PLoS ONE*. 2015;10(9):e0137203. doi: 10.1371/journal.pone.0137203.
21. Scholtens S, Smidt N, Swertz MA, et al. Cohort profile: LifeLines, a three-generation cohort study and biobank. *International Journal of Epidemiology*. 2015;44(4):1172-80. doi: 10.1093/ije/dyu229.
22. Ballering AV, Bonvanie IJ, Olde Hartman TC, Monden R, Rosmalen JGM. Gender and sex independently associate with common somatic symptoms and lifetime prevalence of chronic disease. *Social Science & Medicine*. 2020;253:112968. doi: 10.1016/j.socscimed.2020.112968.
23. Witchel SF. Disorders of sex development. *Best Practice & Research Clinical Obstetrics & Gynaecology*. 2018;48:90-102. doi: 10.1016/j.bpobgyn.2017.11.005.
24. Bem SL. The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*. 1974;42(2):155-162. doi: 10.1037/h0036215.
25. Donnelly K, Twenge JM. Masculine and feminine traits on the bem sex-role inventory, 1993–2012: A cross-temporal meta-analysis. *Sex Roles*. 2017;76(9):556-565. doi: 10.1007/s11199-016-0625-y.
26. Lippa R, Connelly S. Gender diagnosticity: A new bayesian approach to gender-related individual differences. *Journal of Personality and Social Psychology*. 1990;59(5):1051-1065. doi: 10.1037/0022-3514.59.5.1051.
27. Smith PM, Koehoorn M. Measuring gender when you don't have a gender measure: Constructing a gender index using survey data. *International Journal for Equity in Health*. 2016;74(6):459-468. doi: 10.1186/s12939-016-0370-4.
28. Pelletier R, Ditto BF, Pilote L. A composite measure of gender and its association with risk factors in patients with premature acute coronary syndrome. *Psychosomatic medicine*. 2015;77(5):517-526. doi: 10.1097/PSY.0000000000000186.
29. Muilwijk M, Bolijn R, Galenkamp H, Stronks K, van Charante EM, van Valkengoed IG. The association between gender-related characteristics and type 2 diabetes risk in a multi-ethnic population: The HELIUS study. *Nutrition, Metabolism and Cardiovascular Diseases*. 2022;32(1):142-150. doi: 10.1016/j.numecd.2021.09.015.
30. McHale SM, Kim JY, Dotterer AM, Crouter AC, Booth A. The development of gendered interests and personality qualities from middle childhood through adolescence: A biosocial analysis. *Child development*. 2009;80(2):482-495. doi: 10.1111/j.1467-8624.2009.01273.x.
31. Egan SK, Perry DG. Gender identity: A multidimensional analysis with implications for psychosocial adjustment. *Developmental Psychology*. 2001;37(4):451-463. doi: 10.1037/0012-1649.37.4.451.
32. Potter A, Dube S, Allgaier N, et al. Early adolescent gender diversity and mental health in the adolescent brain cognitive development study. *Journal of Child Psychology and Psychiatry*. 2021;62(2):171-179. doi: 10.1111/jcpp.13248.
33. Joel D, Tarrasch R, Berman Z, Mukamel M, Ziv E. Queering gender: Studying gender identity in 'normative' individuals. *Psychology and sexuality*. 2014;5(4):291-321. doi: 10.1080/19419899.2013.830640.
34. Deogracias JJ, Johnson LL, Meyer-Bahlburg H, Kessler SJ, Schober JM, Zucker KJ. The gender identity/gender dysphoria questionnaire for adolescents and adults. *The Journal of Sex Research*. 2007;44(4):370-379. doi: 10.1080/00224490701586730.

35. Singh D, Deogracias JJ, Johnson LL, et al. The gender identity/gender dysphoria questionnaire for adolescents and adults: Further validity evidence. *Journal of Sex Research*. 2010;47(1):49-58. doi: 10.1080/00224490902898728.
36. Bauer GR, Braimoh J, Scheim AI, Dharma C. Transgender-inclusive measures of sex/gender for population surveys: Mixed-methods evaluation and recommendations. *PloS ONE*. 2017;12(5):e0178043. doi: 10.1371/journal.pone.0178043.
37. Lombardi E, Banik S. The utility of the two-step gender measure within trans and cis populations. *Sexuality Research & Social Policy*. 2016;13(3):288-296. doi: 10.1007/s13178-016-0220-6.
38. Savin-Williams RC. Sexual orientation: Categories or continuum? commentary on bailey et al. (2016). *Psychological science in the public interest*. 2016;17(2):37-44. doi: 10.1177/1529100616637618.
39. Brooks H, Llewellyn CD, Nadarzynski T, et al. Sexual orientation disclosure in health care: A systematic review. *The British Journal of General Practice*. 2018;68(668):e187-e196. doi: 10.3399/bjgp18X694841.
40. Cochran SD, Björkenstam C, Mays VM. Sexual orientation differences in functional limitations, disability, and mental health services use: Results from the 2013-2014 national health interview survey. *Journal of consulting and clinical psychology*. 2017;85(12):1111-1121. doi: 10.1037/ccp0000243.
41. Sandfort TGM, Bakker F, Schellevis FG, Vanwesenbeeck I. Sexual orientation and mental and physical health status: Findings from a dutch population survey. *American Journal of Public Health*. 2006;96(6):1119-1125.
42. Guillamon A, Junque C, Gómez-Gil E. A review of the status of brain structure research in transsexualism. *Archives of Sexual Behavior*. 2016;45(7):1615-1648. doi: 10.1007/s10508-016-0768-5.
43. Lawrence AA. Sexual orientation versus age of onset as bases for typologies (subtypes) for gender identity disorder in adolescents and adults. *Archives of Sexual Behavior*. 2010;39(2):514-545. doi: 10.1007/s10508-009-9594-3.
44. Bailey JM, Vasey PL, Diamond LM, Breedlove SM, Vilain E, Epprecht M. Sexual orientation, controversy, and science. *Psychological science in the public interest*. 2016;17(2):45-101. doi: 10.1177/1529100616637616.
45. Salway TJ, Morgan J, Ferlatte O, Hawkins B, Lachowsky NJ, Gilbert M. A systematic review of characteristics of nonprobability community venue samples of sexual minority individuals and associated methods for assessing selection bias. *LGBT health*. 2019;6(5):205-215. doi: 10.1089/lgbt.2018.0241.
46. Hottes TS, Gesink D, Ferlatte O, et al. Concealment of sexual minority identities in interviewer-administered government surveys and its impact on estimates of suicide ideation among bisexual and gay men. *Journal of Bisexuality*. 2016;16(4):427-453. doi: 10.1080/15299716.2016.1225622.
47. Morgenroth T, Sendén MG, Lindqvist A, Renström EA, Ryan MK, Morton TA. Defending the sex/gender binary: The role of gender identification and need for closure. *Social psychological & personality science*. 2020;12(5):731-740. doi: 10.1177/1948550620937188.
48. Suen LW, Lunn MR, Katuzny K, et al. What sexual and gender minority people want researchers to know about sexual orientation and gender identity questions: A qualitative study. *Archives of Sexual Behavior*. 2020;49(7):2301-2318. doi: 10.1007/s10508-020-01810-y.
49. Spiel K, Haimson O, Lottridge D. How to do better with gender on surveys: A guide for HCI researchers. *Interactions*. 2019;26(4):62-65. doi: 10.1145/3338283.
50. Medeiros M, Forest B, Öhberg P. The case for non-binary gender questions in surveys. *Policital Science & Politics*. 2020;53(1):128-135. doi: 10.1017/S1049096519001203.
51. Case P, Austin SB, Hunter DJ, et al. Disclosure of sexual orientation and behavior in the nurses' health study II. *Journal of homosexuality*. 2006;51(1):13-31. doi: 10.1300/J082v51n01_02.

CHAPTER 13

Ballering, A.V., Plug, I., Lagro-Janssen, A.L.M., Das, E., & Rosmalen, J.G.M. (2022). Building a bridge between multidisciplinary insights and practice: The development of an e-learning for internal residents about sex, gender and persistent somatic symptoms. *Journal of Psychosomatic Research*, *155*, 110739.

Abstract

Research findings that may advance healthcare and society are frequently confined to the academic ivory tower, instead of finding their way to healthcare professionals. This is especially the case for multidisciplinary research outcomes. We attempted to build a bridge between multidisciplinary research outcomes and clinical practice. Therefore, we developed an e-learning, aiming to disseminate our research results to clinicians. We aimed to raise awareness among internists in training and their supervisors of the role that sex and gender play in the illness trajectory of people with persistent somatic symptoms, with special attention to the communication between the GP and patient. The e-learning brings together insights into sex, gender and illness trajectories of persistent somatic symptoms from epidemiology, general practice, internal medicine, psychology, sociology, linguistics and communication. We experienced that combining insights from various research domains, translating theoretical knowledge into hands-on clinical tools, and inviting multiple stakeholders relevant to internist training to provide valuable input, is a synergistic approach to achieve this.

Introduction

Research outcomes that may advance healthcare and society are frequently confined to the academic ivory tower, instead of finding their way to healthcare professionals. Particularly knowledge and outcomes from different scientific disciplines are oftentimes not integrated into an interdisciplinary perspective on health, which impedes an evidence-based biopsychosocial approach. We argue that the onus is, at least partly, on the researcher to facilitate the practical and societal implementation of multidisciplinary research outcomes. We developed an e-learning to disseminate our research outcomes to healthcare professionals, following previous initiatives and approaches.[1-3] Our e-learning connects scientific insights into sex, gender, and persistent somatic symptoms (henceforth: PSS) from various disciplines, and functions as a bridge between theory and practice.

This e-learning is built upon acquired empirical and multidisciplinary general knowledge on the importance and implementation of sex and gender in healthcare and in medical curricula.[1,3-5] Specifically, we focus on internal residents, as sex and gender sensitivity and awareness are an obligatory competence during the Dutch education of these specialists.[6] Furthermore, we consider internal medicine as an especially relevant specialty to focus on, since approximately 61% of the patients who attend a general internal outpatient clinic experience PSS.[7]

To clearly understand the relevance of sex and gender in relation to PSS, we should first distinguish between the two concepts. Sex refers to the biological characteristics of female and male bodies, such as genes, hormones and physiology. Gender entails the embodiment of different identities, roles and behaviors of men and women prescribed by societal norms in a given time and society.[8] Patients' sex and gender are independently associated with more frequent and more severe persistent symptoms of women, as well as with physicians' decisions regarding to diagnosis and treatment of PSS in women and men.[9-12] Both sex and gender affect the epidemiology of persistent somatic symptoms. Gender also influences the physician's approach and communication between physicians' and patients in consultations, which is important for physicians' decisions about healthcare.[13] Gender plays a pivotal role in especially the content of physicians' and patients' communication; in other words, in what is actually said.[14,15] Gender differences in the manner of communication, so how something is said, are far less pronounced.[16,17]

To synthesize relevant findings from different scientific disciplines to clinical implications, and to increase awareness of the roles of sex, gender, and communication in healthcare trajectories of patients with PSS among internists, we have developed

an e-learning for internists in training and their supervisors. Connecting valuable insights from various disciplines, our e-learning functions as a bridge between multidisciplinary insights and medical practice. Advantages of an e-learning are the large reach, the easy and flexible accessibility for students, and direct application of the obtained knowledge into practice.[18] Our e-learning was developed with the intention to have internists in training fulfil the following objectives: (1) understanding the differences between women and men in prevalence of and predisposition to PSS, (2) recognizing and being wary of gender bias in communication, and (3) being able to identify gender-related factors that are important to consider when treating patients with PSS.

Contents and development

The fictitious case of a 32-year-old female patient with persistent abdominal pain runs as a thread through the e-learning. The case is introduced in a short referral letter from the general practitioner (GP) to the internist, followed by a short video of the patient's first consultation with the female internist. Hereafter, in the first of three modules, the concepts of sex and gender, as well as sex- and gender-differences in the epidemiology of PSS are introduced and explained. This module combines insights from epidemiology, general practice, psychology, and sociology. Then, the second module assesses unconscious stereotypes of women's and men's communication, and demonstrates how these stereotypes affect gender inferences, and patients' and physicians' communication. This module combines recent findings from linguistics, communication, general practice and sociology. The last module of the e-learning focuses on optimizing internists' professional interaction with patients affected by PSS. These gender-sensitive consultations are characterized by a model that assesses the somatic, cognitive, emotional, behavioral and social aspects in PSS care and consultations.

We involved internists in-training, practicing and supervising internists, GPs, and patient representatives in all stages of the development of the e-learning, to make sure that all objectives and modules' content were understandable, recognizable, and applicable to clinical practice. An institutional learning template was used to ensure an attractive and motivating learning environment. To enable and stimulate (inter-) active learning we included priming questions, mini lectures and transcripts of real doctor-patient interactions to illustrate the impact of gender on PSS and related consults throughout the e-learning. Ultimately, ten questions related to sex, gender, PSS and communication test participants' obtained knowledge. Upon successfully

completing the e-learning a certificate and accreditation are provided to participants. Completion of the e-learning takes approximately 90 minutes. The e-learning was developed in Dutch. We are currently preparing an English translation of this e-learning, to reach more physicians. Moreover, because sex and gender sensitivity are now designated as a compulsory competence in the Dutch training in internal medicine, the e-learning's application will be facilitated in relevant educational programs for internists in training. Additional information regarding the content of the e-learning is available upon request. We aim to add these modules to an existing interprofessional e-learning on PSS to also reach other healthcare disciplines[2].

Conclusion

The e-learning's interactive training modules increase awareness of the importance of sex and gender, in healthcare for patients with PSS, and provide internists with pragmatic tools for applying scientific knowledge into daily practice. We recommend for fellow researchers to move their research beyond the ivory tower by facilitating the implementation of their multidisciplinary research outcomes. Developing an e-learning is merely one of the available ways to build such a bridge between knowledge from various disciplines and medical practice. We experienced that combining insights from various research domains, translating theoretical knowledge into hands-on clinical tools, and inviting multiple stakeholders relevant to internist training to provide valuable input, is a synergistic approach to achieve this.

References

1. Dielissen P, Verdonk P, Waard MW, Bottema B, Lagro-Janssen T. The effect of gender medicine education in GP training: A prospective cohort study. *Perspectives on Medical Education*. 2014;3(5):343-356. doi: 10.1007/s40037-014-0122-3.
2. van Gils A, Tak LM, Sattel H, Rosmalen JGM. Development and user experiences of a biopsychosocial interprofessional online course on persistent somatic symptoms. *Frontiers in psychiatry*. 2021;12:725546. doi: 10.3389/fpsyt.2021.725546.
3. Lagro-Janssen T. Sex, gender and health: Developments in research. *European Journal of Women's Studies*. 2007;14(1):9-20. doi: 10.1177%2F1350506807072314.
4. Scholte JK, van der Meulen FW, Teunissen TA, et al. Exploring the views of successful applicants for medical school about gender medicine using a gender-sensitive video assignment. *BMC medical education*. 2020;20(1):1-9.
5. van der Meulen F, Fluit C, Albers M, Laan R, Lagro-Janssen A. Successfully sustaining sex and gender issues in undergraduate medical education: A case study. *Advances in Health Sciences Education*. 2017;22(5):1057-1070.
6. Lagro-Janssen T, Scholte J, Pepping T. Gendersensitieve zorg in de geneeskundige vervolgopleidingen: Eindrapportage CGS-project. (gendersensitive healthcare in the medical education: Final report of the board of medical specialistic education) . *Radboudumc, Gender & Women's Health*. 2018.
7. Gol JM, Burger H, Janssens KA, Slaets JP, Gans RO, Rosmalen JGM. PROFSS: A screening tool for early identification of functional somatic symptoms. *Journal of Psychosomatic Research*. 2014;77(6):504-509. doi: 10.1016/j.jpsychores.2014.10.004.
8. Johnson JL, Greaves L, Repta R. Better science with sex and gender: Facilitating the use of a sex and gender-based analysis in health research. *International Journal for Equity in Health*. 2009;8(1):14. doi: 10.1186/1475-9276-8-14.
9. Ballering AV, Wardenaar KJ, olde Hartman TC, Rosmalen JGM. Female sex and femininity independently associate with common somatic symptom trajectories. *Psychological Medicine*. 2022;52(11):2144-2154. doi: 10.1017/S0033291720004043.
10. Ballering AV, Muijres D, Uijen AA, Rosmalen JGM, olde Hartman TC. Sex differences in the trajectories to diagnosis of patients presenting with common somatic symptoms in primary care: An observational cohort study. *Journal of Psychosomatic Research*. 2021;149:110589. doi: 10.1016/j.jpsychores.2021.110589.
11. Ballering AV, Bonvanie IJ, Olde Hartman TC, Monden R, Rosmalen JGM. Gender and sex independently associate with common somatic symptoms and lifetime prevalence of chronic disease. *Social Science & Medicine*. 2020;253:112968. doi: 10.1016/j.socscimed.2020.112968.
12. Groeneveld JM, Ballering AV, van Boven K, Akkermans RP, Olde Hartman TC, Uijen AA. Sex differences in incidence of respiratory symptoms and management by general practitioners. *Family practice*. 2020;37(5):631-636. doi: 10.1093/fampra/cmaa040.
13. Houwen J, Lucassen P, Stappers HW, Assendelft PJJ, van Dulmen S, Olde Hartman TC. Medically unexplained symptoms: The person, the symptoms and the dialogue. *Family Practice*. 2017;34(2):245-251. doi: 10.1093/fampra/cmw132.
14. Hall JA, Irish JT, Roter DL, Ehrlich CM, Miller LH. Gender in medical encounters: An analysis of physician and patient communication in a primary care setting. *Health Psychology*. 1994;13(5):384. doi: 10.1037//0278-6133.13.5.384.

15. Roter DL, Hall JA, Aoki Y. Physician gender effects in medical communication: A meta-analytic review. *JAMA*. 2002;288(6):756-764. doi: 10.1001/jama.288.6.756.
16. Aries E. *Men and women in interaction: Reconsidering the differences.* Oxford University Press; 1996.
17. Plug I, Stommel W, Lucassen P, Olde Hartman T, Van Dulmen S, Das E. Do women and men use language differently in spoken face-to-face interaction? A scoping review. *Review of Communication Research*. 2021;9:43-79. doi: 10.12840/ISSN.2255-4165.026.
18. Rosmalen JGM, van Gils A, Acevedo Mesa MA, Schoevers RA, Monden R, Hanssen DJ. Development of grip self-help: An online patient-tailored self-help intervention for functional somatic symptoms in primary care. *Internet interventions*. 2020;19:100297. doi: 10.1016/j.invent.2019.100297.

CHAPTER 14

General discussion

The aim of this thesis was to gain in-depth insights into whether and how sex and gender are associated with the illness trajectories of common somatic symptoms, while taking into account the multifaceted nature of sex and gender. Although previous epidemiological studies have indicated that sex differences occur at multiple stages of illness trajectories of various health conditions, these studies apply oversimplified methodologies and are limited in their design and interpretation. In this thesis, insights into sex and/or gender differences in illness trajectories were generated from an interdisciplinary stance, allowing for an integration of epidemiology, (bio) medical science, psychology, ethics, pedagogy and sociology.

This final chapter of the thesis draws up the balance. It first summarizes the included studies' main findings. Second, methodological considerations will be discussed. Third, three themes that emerged from this thesis, namely (1) gender in epidemiology; (2) rethinking sex and gender differences in illness trajectories; and (3) the future of sex and gender sensitive medicine (SGSM), will be discussed in light of previous scientific literature and societal developments. Fourth, implications for policy, clinical practice and for future research are discussed. Last, an overarching conclusion is provided.

Main findings

The main findings of this thesis are described in the following paragraph, structured by an illness trajectory that may be initiated for common somatic symptoms. These illness trajectories start when a person notices a bodily sensation that is interpreted to be "wrong" and only ends if the symptom is resolved and the provided care for the symptom is considered to be finished. During an illness trajectory, which may last from a mere few hours to many years, multiple critical junctions occur. These are turning points involving an experience or event, or interdependent sequence of events, that have a potentially far-fetching impact on the patient's health and healthcare experience. These critical junctions include, but are not limited to the interpretation of the severity of symptoms, seeking healthcare for symptoms and being provided with a diagnosis for symptoms. It should be stressed that being provided with a diagnosis for somatic symptoms may not necessarily be the end of an illness trajectory. Although illness trajectories are complex and iterative, with many twists and turns, for the sake of simplicity we describe the main findings structured by a linear illness trajectory.

This thesis starts by exploring the first critical junction in an illness trajectory, namely the interpretation of a bodily sensation as a symptom. To this end, we study the biological and psychosocial factors in which the development and persistence of somatic symptoms are rooted. Women and men differ herein. The first section of this thesis shows that sex differences occur in biological factors contributing to common somatic symptoms. We aimed to replicate a small-scale study that reported a significant association between the rs9470080 genotype and common somatic symptom levels in women, but not in men (*Chapter 2*).[1] We could not replicate these results in the large-scale general population cohort study Lifelines, as the rs9470080 genotype did not associate with somatic symptoms in neither male or female participants. The overall genetic contribution to somatic symptoms was found to be higher in men than in women. Not only genetic factors, but also childhood experiences may affect one's proneness to common somatic symptoms. In *Chapter 3* we used data derived from the prospective Tracking Adolescents' Individual Lives (TRAILS) study.[2] We found that parents of adolescent boys reported more symptoms than the boys themselves, while parents of adolescent girls reported fewer symptoms than the girls themselves. The level of parental-reported somatic symptoms during adolescence associated significantly with somatic symptoms experience in childhood, but the strength of this association did not differ over levels of adolescents' self-report or sex.

In adulthood, biological and psychosocial factors remain important in common somatic symptoms, as we show in the second section of this thesis. In *Chapter 4* we examined the cross-sectional, independent associations between sex, gender, and the prevalence of common somatic symptoms.[3] This study was based on Lifelines data. We describe a data-driven method to develop a gender measure (i.e., the gender index) in a cohort that did not include items to directly measure any dimension of participants' gender. We found that both female sex and feminine gender, operationalized by the gender index, associated with common somatic symptoms and chronic disease. Feminine gender associated more strongly with common somatic symptoms in male participants than in female participants. Subsequently we studied the independent associations between sex, gender, and the somatic symptom severity in a longitudinal study (*Chapter 5*).[4] We identified five linear trajectories of common somatic symptom severity over time in a data-driven manner. These varied from a stable and low severity trajectory in the vast majority of the population, to a trajectory with increasing symptom severity in relatively few participants. When we compared these participants, we found that female sex positively associated with a higher symptom severity, whereas feminine gender negatively associated with this. In *Chapter 6*, we also describe the severity of 23 somatic symptoms surrounding a COVID-19 diagnosis.[5] Due to the unique nature of the Lifelines COVID-19 Cohort Study

we could account for individuals' symptom severity before the COVID-19 diagnosis, as well as symptom severity among an uninfected control population when assessing symptom persistence. We identified three types of symptoms: (1) acute symptoms in which symptom severity was worsened in the days surrounding the COVID-19 diagnosis, but returned to pre-COVID-19 levels within 50 days; (2) core symptoms in which symptom severity persisted and did not return to pre-COVID-19 levels within 50 days and (3) other symptoms, for which no distinctive pattern was observed. Visual inspection of the data showed that during the acute phase of COVID-19, women reported higher severity for the majority of acute symptoms and core symptoms. The core symptoms also persisted longer with a higher severity level in women compared to men. We considered both female sex and feminine gender to be important in this female preponderance in symptom experience, but did not test this statistically.

Another critical junction in the illness trajectory for common somatic symptoms is the decision to seek help from a healthcare professional for your symptoms. We studied potential sex and gender differences in this critical junction in the third section of this thesis (*Chapter 7*).[6] To this end we linked data from the general population cohort Lifelines with the Nivel Primary Care Database. We show that female sex rather than feminine gender, operationalized by the gender index, was associated with consulting the general practitioner (GP) for common somatic symptoms. Patients' number of paid working days was negatively associated with primary care help-seeking for common somatic symptoms. Previous research frequently attributed sex differences in help-seeking behavior to gender differences between women and men. Our results suggest that factors related to the frequency of help-seeking are rooted in biology, or in components that are beyond the composite gender index. In our view, the gendered social frames imposed upon women and men by society, complemented by biological factors, shape help-seeking behavior.

After the decision to seek help is made, another critical junction takes place in the consulting room of the GP, namely whether the patient is provided with diagnostic interventions. The thesis' fourth section examined the sex differences herein. We show in *Chapter 8* and *Chapter 9* that male patients who presented themselves with cough, dyspnea or other common somatic symptoms at the GP, were provided with more diagnostic interventions, such as a physical examination, diagnostic imaging, and a referral to a specialist by the GP than female patients.[7,8] In contrast, female patients with these symptoms were more frequently provided with laboratory diagnostics than male patients. These sex differences in provided diagnostic interventions contribute to the 6%-lower odds that women have of their symptoms being diagnosed with a disease compared to men. We argue that many factors, including biological processes, gendered

stereotypes, communication, and GPs' potential diagnostic uncertainty, interact and partially explain the sex differences in provided interventions. In *Chapter 10* we show that irrespective of their frequency of being provided, diagnostic interventions are less effective in contributing to a disease diagnosis in female patients than in male patients.[9] This may also partly explain why GPs provide fewer diagnostic interventions to female patients than to male patients. Similarly, in the COVID-19 pandemic sex and gender differences in diagnostics were observed as well (*Chapter 11*).[10,11] Within the population of healthcare workers, women were less frequently diagnosed and tested than men. This male preponderance in testing and diagnosis could be explained by the more pronounced COVID-19 symptoms in men, and worse prognosis of COVID-19, compared to women (*Chapter 11A*). Institutionalized gender inequities may play a role herein.

In the last section of this thesis we reflected on sex and gender in epidemiological studies and we described the pitfalls we encountered when conducting sex and/or gender-focused epidemiological research. We also describe how we overcame these caveats.[12] In addition, we provide recommendations for the future assessment and inclusion of sex and gender measures in large-scale cohort studies (*Chapter 12*). Last, the development of a new e-learning approach to disseminate knowledge about sex and/or gender differences in illness trajectories of people with common somatic symptoms among key target groups is described (*Chapter 13*).[13]

In summary, this thesis provides an extensive overview of how sex and gender associate with the illness trajectories of common somatic symptoms. We show that many critical junctions in an illness trajectory, including the interpretation of a symptom's severity, help-seeking behavior and diagnostics, are affected by sex and gender, resulting in different health outcomes for men and women.

Methodological considerations

When interpreting the results of the studies presented in this thesis, we should acknowledge several methodological limitations. Although specific methodological considerations have been discussed in-depth in the corresponding chapters, we concisely reflect on some important general considerations that we feel deserve further attention.

First, multiple studies in this thesis (*Chapter 4, 5, and 7*) made use of the data-driven gender index to operationalize gender. Although the advantages and disadvantages of the methodology underlying the gender index will be described in detail below, it

should be noted that the use of a gender index in itself is by no means a replacement for validated self-reported measures that assess the embodiment of the different dimensions of gender (*Chapter 12)*. A gender index functions as a proxy measure for (dimensions of) gender, but does not capture gender in all its facets. Individuals' embodiment of gender may change over time, as well as what is considered feminine, masculine, neither or androgenous (i.e., a combination of masculinity and femininity). As the results in this thesis are based on a gender index that was developed based on Lifelines' baseline data from 2006 to 2012, it is likely that the composition of the gender index would slightly change if it were to be developed by using the most recent Lifelines data. Therefore, inherent to the dynamic nature of gender, the generalizability of results obtained by the gender index we developed is limited and bound to the context in which the research took place. The methodology underlying the gender index is nevertheless adaptive and applicable to large-scale cohort studies in all types of settings.

Second, many empirical studies in thesis used data derived from large observational general population cohorts, such as Lifelines (*Chapters 2, 4* to *6,* and *11*) and TRAILS (*Chapter 3*). *Chapter 7* combined Lifelines data with the Nivel Primary Care Registration Database. When using data from large-scale observational cohorts, a risk of selection bias is present. Potentially only participants who are intrinsically interested in health agreed to be included in the study. Yet, the Lifelines Cohort Study and TRAILS were previously shown to be largely representative of the general population.[14,15] Although these large observational general population cohort studies provide us with a wealth of data, due to an already significant participant burden not all health-related details, such as frequency and reasons for healthcare seeking, can be assessed in detail. To a certain extent this can be countered by linking data as was done in *Chapter 7*, which allows for the combination of two large datasets with rich data. This linkage enabled us to assess sex and gender-differences in the frequency of healthcare seeking, but the linked data did not include information on reasons for encounter. FaMe-Net does include the latter information, but in turn does not suffice for developing a data-driven gender index. Although general population cohort studies, such as Lifelines and TRAILS provide valuable insights into public health, it needs to be kept in mind that research results obtained from these cannot be directly translated to clinical populations. In contrast, the studies based on FaMe-Net data are conducted in a primary care patient population that is representative for the general population in terms of age and sex,[16] allowing for more straightforward recommendations for policy and clinical practice. The complementarity of these cohorts results in a more nuanced perspective on sex and/or gender differences in illness trajectories for common somatic symptoms.

Third, this thesis focusses on sex and/or gender differences in illness trajectories of 12 common somatic symptoms. These symptoms are derived from the validated Symptom CheckList-90 Somatization (i.e., SCL-90 SOM) subscale.[17] However, the adjective 'common' is applicable to different symptoms in different settings. As can be derived from *Chapter 7*, new-onset heart pain and shortness of breath are only reported by 2.7% and 3.1% of the general population. Within the primary care population, chills and muscle pain are the least frequent reasons for encounter by 0.3% and 1.5% of the patient population, respectively. Episodes of care initiated with general tiredness or lower back pain, in contrast, comprise over one-third of the total number of episodes of care based on the twelve symptoms in the SCL-90. This thesis shows that the commonness of symptoms is not similar across settings. This should be kept in mind when interpreting the research results, as other characteristics of symptoms (e.g., their consequences for quality of life) may have a larger influence on whether patients seek help than symptoms' commonness.

Fourth, *Chapters 8* to *10* are based on FaMe-Net data and dichotomize disease diagnoses (i.e., symptoms that, followed over time, can be attributed to a disease, operationalized as ICPC≥70) and symptom diagnoses (i.e., symptoms, followed over time, for which the relevant diagnostic criteria of a disease are not fulfilled and that cannot be attributed to a pathophysiological disease or syndrome, operationalized as ICPC<30).[7-9,18] Symptom diagnoses are roughly considered to be similar to medically unexplained symptoms or functional somatic symptoms. These symptoms are defined as symptoms that persist for several weeks for which no sufficient medical explanation (i.e., disease or bodily abnormality) can be found despite adequate medical history-taking and diagnostic testing.[19] Although the validity of the FaMe-Net data registration is relatively high due to the structural peer-to-peer reflection to evaluate diagnostic criteria to minimize misclassification during registration, it remains challenging to uniformly distinguish whether somatic symptoms are sufficiently or insufficiently explained by psychiatric and/or somatic diseases. Such a decision may depend on the patient's medical history and clinical presentation, and the clinical gaze of the GP. The diagnostic process is complicated by discrepancies in the degree of objective assessments of bodily dysfunction or pathology and the subjective symptoms people may experience due to these. Furthermore, the extent with which symptoms are diagnosed as a symptom diagnoses may vary between GPs, due to differing clinical opinions and experience. The uncertainty in whether somatic symptoms are sufficiently explained may introduce noise in the data registration and potentially stimulate a bias towards symptom diagnoses in women, as research shows that the severity

of women's symptoms tend to be underestimated.[20,21] Although one may argue that self-reported measures completed by patients regarding the nature of their symptoms (i.e., roughly dichotomized into functional or non-functional) would be more precise, this measure is far from perfect as well due to recall bias and lack of medical knowledge among patients.

Last, although the chapters in this thesis mainly rely on quantitative data, we tried to incorporate the lived experience of people with common somatic symptoms as a guiding principle throughout our research. To this end, we set up a patient panel that was consulted at several points in the research project. We aimed to create a diverse and inclusive patient panel in which panelists differed in, among others, their sex, gender identity, age and sexual orientation. We experienced difficulties in setting up this panel, regarding diverse inclusion. We could have countered this by collaborating more closely with healthcare organizations that work more frequently with transgender and gender diverse patients. Furthermore, no patient association for common somatic symptoms exist. We also faced challenges in the collaboration with the patient panel. A hierarchy between the patient panel and us as researchers occurred, even though we intended to avoid this by providing the panelists with a short course on science and the scientific process. This probably made it hard for panelists to openly discuss their experiences and ideas. Additionally, we as researchers were not fully sure on how to make use of the experiences and knowledge of the patient panel and only consulted the panel sporadically. The incidental nature of these consultations, with us as scientists requiring input on a specific challenge, may have further increased the hesitance of panelists to openly express themselves on topics they considered important. Therefore, the studies in this thesis could have benefitted more from patients' guidance. This experience has taught us that collaborating with patients or non-researchers in studies requires a skillset that should be trained and cultivated.

Emerging themes

1. Gender in epidemiology

This thesis underlines the relevance of gender, in addition to sex, in epidemiology. Traditionally gender did not have a prominent place in epidemiological studies and was considered to be a mere consequence of, and therefore similar to, people's sex assigned at birth.[12] Only after its independence from sex and its relevance to health became more prominent in mainly sociological research, a variety of quantitative methodologies were developed and implemented to assess gender in epidemiological studies.[22]

Before the introduction of the Bem Sex Role Inventory (BSRI) in 1974, femininity and masculinity were considered as contrasts in Western societies.[22] A substantial body of evidence has broken down the then-existing dogma of opposing femininity and masculinity. The BSRI introduced a complementary way of assessing masculinity and femininity by acknowledging their co-occurrence in individuals.[22-24] The BSRI is one of the first and most widely-used instruments that uses this bidimensional, quantitative approach of gender.[25] Despite criticism towards the BSRI and although one would expect that due to the time, place, and culture-sensitive nature of gender the BSRI is abandoned as a valid measure for gender, it is still frequently used in healthcare research, albeit sometimes adapted.[26] It is either employed as a stand-alone measure for gender,[24] but also as part of more comprehensive gender measures.[27]

The renewed attention towards gender and health also sparked an interest in the development of comprehensive methodologies that assess gender in epidemiological studies that lack gender measures in the initial data collection. Especially gender indices have gained increased attention in the last decade. Herein, multiple combined components define masculinity and/or femininity usually in terms of gender roles. A variety of gender indices have been developed in roughly the past decade, mainly via fully theory-driven methodologies or via a combination of theoretical and conventional statistical approaches (***Table 1***).[27-32] The gender index we describe in *Chapter 4* is the only gender index known to us that is based on a fully data-driven, machine-learning methodology.[3]

Recently, the methodologies underlying gender indices, and whether these should be data-driven or theory-driven, have been questioned.[33] Theory-driven gender indices are limited in their utility for a variety of reasons. First, some of the developed gender indices do not take into account the broadness and multidimensionality gender, but merely focus on one domain. An example is the unidimensional Labor Force Gender Index (LFGI) that focusses solely on people's gender roles and institutionalized gender in relation to their occupation, disregarding other important domains such as leisure activities, lifestyle and (social) mobility.[32] Second, once developed, the contents of a theory-driven gender index are static, while the embodiment of gender roles is an ever-changing process. Therefore, a constant redefinition of what components define femininity and masculinity is required in a theory-driven gender index. The recently developed one-dimensional masculine gender score in the Dutch Doetinchem Cohort Study includes variables on education, with an educational level higher than one's partner indicating masculinity.[28] Currently, in the Netherlands the proportion of women with a high

educational level has been steadily increasing, surpassing that of men in the early 2000's.[34] This shows how components defining masculinity may evolve over time. Third, theory-driven gender indices rely heavily on expert knowledge. However, experts are not free from bias, potentially reinforcing sexism in the development of gender indices. Fourth, the components of fully theory-driven gender indices frequently have an equal weight in defining femininity and/or masculinity, while these may differ in their extent of contributing to femininity and masculinity. This may result in imprecise operationalizations of gender.

Despite the aforementioned limitations of theory-driven gender indices, a fully data-driven gender index, as defined in *Chapter 4* of this thesis, is not necessarily preferable in all study designs. The methodology underlying this gender index is versatile, flexible and applicable to many large-scale cohort studies due to its capability to account for the time, place, and culture-sensitivity of gender roles, yet the gender index is defined as a unidimensional, bipolar scale.[3] Herein the underlying algorithm forces psychosocial components to be either predicting female or male sex, reinforcing mutual exclusivity. As described in *Chapter 12,* participants are subsequently forced to have a feminine, masculine, or androgenous score on the gender index.[12] This disregards the possibility for the gender index to indicate neither femininity nor masculinity. It is difficult to counter this, as generally current (partly) data-driven gender indices are derived from a form of logistic regression analyses. Herein the outcome (i.e., participants' sex) is dichotomous with the inverse of male sex automatically being female sex. In addition, large datasets are required that allow for sufficient variance of the included predictors to calculate a gender index per participant via the aforementioned methodology. Due to this reason, we were unable to develop a gender index in the studies using FaMe-Net data (*Chapters 8, 9, and 10*).[7-9] Within FaMe-Net few relevant variables were available that could sufficiently explain the variance in sex. The availability of a large variety of variables in a cohort allows for testing many combinations of factors that most optimally associate with female or male sex. This is advantageous for the validity of the analyses. The availability of many variables is only beneficial if the collected data hereon is of high quality, as a data-driven gender index is only as good as the dataset on which the algorithm is trained.

Table 1. Overview of characteristics of composite gender indices.

	Authors (year)	Methodology	Cohort
1	Vader *et al.* (2023)[28]	Theory-driven (one-dimensional)	Doetinchem Cohort Study (2008-2012; N=4,017; 53% female participants)
2	Koehoorn and Smith (2016)[32]	Theory-driven	Canadian Labour Force Survey (1997; N=696,350; 2014; N=729,132; proportion of female participants not reported)
3[a]	Lacasse *et al.* (2020)[31]	Combination of theory-driven and data-driven	Canadian Community Health Survey (2007-2012; N=29,470; 47% female participants)
4	Levinsson *et al.* (2022)[30]	Combination of theory-driven and data-driven	UK Biobank (2006-2010; N=315,937; 53% female participants)

Gender measure	Included components
Masculine gender score, ranging from 0-19 with higher scores indicating more masculine traits	(1) Work and education a. Division of paid work between respondent and partner b. Physical intensity of work c. Educational level compared to partner (2) Informal care a. Time spent on household chores b. Time spent on odd jobs c. Frequency of taking care of sick people (3) Lifestyle a. Physical intensity/type of sport b. Smoking cigars or pipe c. Type of alcohol consumption (4) Emotions a. Limitations in work and activities due to emotional problems b. Experiencing feelings of nervousness c. Feeling energetic and vibrant d. Feeling exhausted and tired
Labour Force Gender Index (LFGI), ranging from 0-10 with lower scores indicating masculine labor market gender roles	(1) Responsibility for caring for children a. Level of reduction in labor market participation due to family responsibilities (2) Occupational segregation a. Male-dominated occupation (3) Hours of work relative to partner/spouse (4) Education relative to partner/spouse
GENDER index, ranging from 0-100 with higher scores indicating having more feminine characteristics	(1) Occupation and education (2) Household composition and income (3) Racial/cultural group (4) Ownership of the household (5) Sense of belonging to the local community (6) Frequency of experienced stress
Femininity Score, standardized in the general population, expressed in standard deviations	(1) Education (2) Occupational status (3) Depression (4) Risk taking (5) Neuroticism (6) Birthyear

Table 1. Continued.

	Authors (year)	Methodology	Cohort
5	Nauman *et al.* (2021)[29]	Combination of theory-driven and data-driven	BASE-II (2009-2014; N=1,869; 51% female participants)
6	Pelletier *et al.* (2015)[27]	Combination of theory-driven and data-driven	GENESIS-PRAXY (2009-2013; N=1,075; 32% female participants)
7[b]	Ballering *et al.* (2020)[3]	Data-driven (machine learning)	Dutch Lifelines Cohort Study (2006-2014; N=152,728; 59% female participants)
8	Lippa and Connely (1990)[35]	Data-driven	Psychology students (Period of data collection not reported; N=227; 48% female participants)

[a]: As 19 components were included in the GENDER Index, we summarized for reasons of clarity. A full overview of the included components is provided in the original study.

Gender measure	Included components
Gender score, ranging from 0-100 with higher scores indicating having more feminine characteristics	(1) Chronic stress (2) Marital status (3) Risk taking behavior (4) Agreeableness (5) Neuroticism (6) Extraversion (7) Loneliness (8) Conscientiousness (9) Education
Gender score, ranging from 0-100 with higher scores indicating having more feminine gender-related characteristics	(1) Primary earner status (2) Personal income (3) Number of hours per week doing housework (4) Primary responsibility for doing housework (5) Level of stress at home (6) Bem Sex Role Inventory – masculinity score (7) Bem Sex Role Inventory – femininity score
Gender index, ranging from 0-100 with higher scores indicating having more feminine characteristics	(1) Type of leisure activities (2) Occupation-related components (3) Time spend on household tasks (4) Time spend on odd jobs (5) Lifestyle (6) Experiencing long-term difficulties or negative life events (7) Personality traits and emotions
Gender diagnosticity measure, ranging from 0-100 with higher scores indicating having more masculine characteristics	(1) Occupational preference

[b]: As 153 (dummy) variables representing 85 unique variables were included in the Gender index, we grouped for reasons of clarity. A full overview of the included components is provided in the original study.

Data-driven gender indices in which sex is regressed on a large variety of psychosocial variables that aim to capture gender, rely on the fundamental notion of gendered characteristics being derived from sex assigned at birth. In other words, within the gender index as defined in *Chapter 4,* femininity or masculinity is defined based on which psychosocial variables are able to significantly differentiate between sexes.[3] This also touches upon one of the main critiques towards data-driven gender indices: a potentially unjustified belief that a data-driven gender index is fully independent from sex.[33] However, when assessing the associations between sex, the gender index as developed in *Chapter 4*, and common somatic symptoms, there were no indications for multicollinearity present. This implies that the statistical model was able to disentangle sex and gender and that these concepts were two statistically separate entities. A second indication of the independence of sex and data-driven gender indices is the great variability in gender scores in both female and male participants.[3,28,30,31,35] Notably, one's beliefs regarding the relationship between sex and gender may influence the appropriateness of defining a data-driven gender index. If one argues sex and gender to be in a continuous dialogue, shaping each other over time, it is implied that a fully data-driven gender index in which sex is predicted by psychosocial variables is inadequate to capture gender, since their association is simultaneous. This indicates the existence of a grey area in which sex and gender are too strongly intertwined, rendering a separation of the two concepts a mere statistical and arbitrary exercise. Techniques, such as directed acyclical graphs, in which the direction of an association between a variety of variables can be identified may be necessary to solve this. However, one could question the clinical relevance of disentangling sex and gender in such a highly detailed manner.

The debate that aims to define the superior end of the methodological spectrum of gender indices is merely theoretical, as in practice nearly similar components define femininity and masculinity in (partly) theory-driven indices[27-32] and the fully data-driven index (***Table 1***).[3] It is especially worth mentioning that the two Dutch indices, namely the aforementioned fully theory-driven masculine gender score and the fully data-driven gender index,[3,28] include highly similar components to define femininity and masculinity. Remarkably, the former study merely compares its masculinity index with previous indices developed in international studies, instead of comparing it with our gender index which was developed in a Dutch population in the same timeframe. Rather than debating about the most adequate one-size-fits-all methodology for gender indices, the applied methodology should be compatible with the study design, research question and sample size. A high-quality, large sample size could warrant the development of a data-driven gender index as such samples allow for (1) including many gendered components; (2) moving beyond the equal-weight components

limitation of theory-driven indices, and (3) training of the underlying algorithm on high quantities of rich data. In contrast, a small sample size or research question that focusses on a specific gender-related component would call for a theory-driven gender index.

The question remains whether gender indices are the most adequate and valid gender measure in epidemiological research. Gender indices are useful tools to obtain a gender measure if no information hereon is collected, but should not be treated as an absolute truth. Components on which gender indices are based may not be gender-sensitively collected. An inherent gender-bias could occur in the construction of survey items or in data collection.[36,37] The former is exemplified by survey items that reproduce hegemonic ideas regarding the distribution of power between men and women, implying superiority of one gender over another.[36] Agreeableness with the statement 'If a woman earns more money than her husband, it's almost certain to cause problems' was, for example, assessed in seventh wave of the large-scale World Value Survey that collected data between 2017-2021 from 64 countries including the Netherlands (N=94,278).[38] A self-perpetuating gender-bias in data collection is illustrated by survey items that assess domains that are stereotypically considered as explicitly female (e.g., children's health) or male (e.g., tobacco use) and are therefore not assessed in the opposite sex.[36] These biases are bound to affect potential components included in gender indices. Furthermore, although specific components in a gender index may strongly associate with health outcomes, these may be obscured when combined with other gendered components into a comprehensive gender index, as shown in *Chapter 7*.[6]

Arguably, rather than combining multiple components into a gender index, it would be more fruitful to include (a combination of) multiple self-reported gender characteristics that represent different domains of gender in statistical analyses as discussed in *Chapter 12*.[12,39,40] Ultimately, we should overcome the need to use gender indices and favor the incorporation of direct gender measures in which participants self-report on the embodiment of dimensions of gender in epidemiological studies. Preferably, multiple assessments over time of gender measures are conducted to capture potential changes in the embodiment of gender over time.[12]

2. Beyond stereotypes: rethinking gender inequity in diagnoses for common somatic symptoms

In the Dutch media it is frequently stated that 80% of people with a symptom diagnosis are women.[41] A symptom diagnosis is provided in case of symptoms for which the relevant diagnostic criteria of a disease are not fulfilled and the symptoms cannot be attributed to a disease. Via anecdotal evidence the 80%-statement has

gained traction over time, is reproduced and is now often assumed to be an absolute truth. However, no recent valid Dutch scientific source underlies this statement. A recent study using FaMe-Net data shows that 63.8% of the patients provided with a persistent symptom diagnosis (i.e., the episode of care lasted for over 1 year) at the GP were women, compared to 57.0% of the patients with a non-persistent symptom diagnosis.[16,42] In addition, *Chapter 9* shows that episodes of care initiated by female patients with common somatic symptoms have 6.3%-higher odds of being concluded with a symptom diagnosis compared to episodes of care initiated by male patients.[7] Although the units of analyses in the latter study are not individuals with or without a diagnosis, but rather episodes of care, and it only focuses on twelve common somatic symptoms instead of all reasons for encounters, the scientific evidence does not indicate an 80%-female preponderance in symptom diagnoses.

With the aforementioned statement being considered as reality, it is often interpreted as an inequity between women and men. We argue that differences between women and men in healthcare do not automatically imply an inequity. *Chapter 9* shows that men are provided with more physical examinations, diagnostic imaging and referrals to a specialist by the GP than women. Women, in contrast, are more frequently provided with laboratory diagnostics for their common somatic symptoms than men. These differences in provided diagnostic interventions between male and female patients, contribute to the lower rates of disease diagnosis in women compared to men.[7]

Multiple factors have been suggested to associate with the difference in provided diagnostic interventions between female and male patients. For example, GPs may experience more diagnostic uncertainty in female patients than in male patients.[43] It has been previously shown that laboratory diagnostics are applied as a first strategy to mitigate diagnostic uncertainty and to avoid anticipated regret of missing a serious disease by GPs, especially in cardiovascular symptoms.[43-45] This may explain the female preponderance in laboratory diagnostics. Another frequently discussed factor that may associate with sex differences in the provided diagnostic interventions is patient-GP communication. A recent scoping review reveals that little robust evidence exists that women and men differ in their language use in one-on-one interactions.[46] Nevertheless, the topics discussed during and the communication style in medical interactions do differ between male and female patients: with female patients more psychosocial topics were discussed than with male patients.[47,48] Male patients are thought to be more dominant in their communication than female patients, by expressing demands for diagnostic strategies.[48] A study that appeared after the aforementioned review shows that female patients do intrusively interrupt their GP more frequently than male patients, which may be interpreted as the result of

female patients feeling the need to make themselves heard and understood in medical interactions.[49] Even when the language use of men and women is the same, gendered preconceptions regarding language and communication are likely to shape people's interpretation of one's language use.[50,51] It is hypothesized that the interpretation of patients' language use by GPs may shape their subsequent medical decision-making, including the provision of diagnostic interventions, as well.[52]

The observed sex difference in diagnostic interventions provided by the GP to patients with common somatic symptoms could be regarded as an inequity that calls for an increase in diagnostic interventions provided to women. After all, if the rate of diagnostic interventions in women would be increased, women would have increased rates of disease diagnosis as well. *Chapter 10*, however, shows that when women are provided with diagnostic interventions, their odds of receiving a disease diagnosis are lower than those of men.[9] This may explain the decreased rate of diagnostic interventions in women, as GPs' clinical experience may have taught them that diagnostic interventions are less effective in women. It has been hypothesized as well that women seek help earlier in their illness trajectory,[53,54] potentially rendering their symptoms more diffuse and more difficult to diagnose. Another reason for the lower effectivity of diagnostic interventions in women is thought to associate with the male-dominated medical research of the past decades, in which women were underrepresented in medical research at all stages.[55] The results presented in *Chapter 10* imply that an equal rate of diagnostic interventions in women and men would not fully resolve the difference in diagnosed disease between women and men. Arguably, we should not require equal rates of disease diagnosis between men and women, as this implies that a disease diagnosis automatically indicates an increased quality of care. Merely arguing for an equal rate in disease diagnoses among men and women, also disregards the consideration that men are overdiagnosed. Rather than only changing the diagnostic behavior of GPs in terms of frequency of diagnostic interventions, efforts should be directed towards developing and validating more sex-sensitive diagnostic interventions and sex-specific guidelines that result in high diagnostic accuracy in both women and men.

The preceding paragraphs highlight two important points. First that assumed truths regarding sex differences may be unsubstantiated by science, and second, that differences in provided healthcare between women and men may be justified. This ties closely to the reification and reproduction of gendered stereotypes in scientific works that allows these stereotypes to gain authority.[56] The adoption of gendered stereotypes within society and their acceptance as reality forms fertile ground for research that automatically assumes that feminine gender or female sex are disadvantaging characteristics. This notion is subsequently further reinforced by such research. Such a process may be analogous to what

previously happened to low socio-economic status (SES). People with low SES were found to be constructed and reified as a problematic group in research and policy documents.[56] People with low SES were, among others, generalized into one homogeneous group, problematized as being an inherently unhealthy population, and characterized as their low SES being a negative personality trait causative of disease. A lack of deconstructing the concept of low SES in research and policy reinforced its status as inherently unhealthy. Parallel mechanisms may be at play when conducting research into sex inequities in health. If feminine gender and female sex are institutionalized as inherently problematic characteristics in research or policy, the responsibility for women's disadvantaged health compared to men could potentially shift towards women instead of remaining at a societal level. This shift in responsibility would be unjust and unfair. Therefore, a critical reflection on how (epidemiological) scientific literature constructs feminine gender and female sex is a prerequisite to avoid misguided policy and research recommendations. This not only refers to simply identifying a conflation of sex and gender. It also involves appraising whether certain (negative) characteristics are attributed to feminine gender or female sex without adequate evidence, or whether studies merely focus on a difference between women and men, overlooking heterogeneity and intersecting social identities within these populations. To avoid misguided recommendations, researchers and policy makers should be critical of research that frames female sex and feminine gender as negative personal characteristics that are causative of disease. Researchers should also be cautious of conducting research that does not take into account the multifaceted nature of sex and gender.

3. The future of sex and gender sensitive medicine: moving beyond sex and gender towards intersectionality

Throughout history, sex and gender sensitive medicine (SGSM) was regarded as relevant to women and it was frequently only performed by women.[57] Although this connotation persists, change is coming as the realization is dawning that SGSM aims to improve health and healthcare for all. The current momentum should be used to routinely integrate sex and gender sensitivity in all branches of medicine. Potentially SGSM could even be used to move beyond merely integrating patients' sex and gender, but additional social identities as well to achieve patient-centric, personalized medicine.

SGSM aims for an inclusive and more precise or individual approach in medicine.[58] Cardiology is a frontrunner in incorporating SGSM, but other medical specialties are lagging behind.[55] For optimal healthcare for all, SGSM should be broadly included in all medical specialties, including those beyond cardiology. Sex and gender sensitivity should become more routinely and deeply ingrained in medicine.[59] This is not limited

to merely an awareness of sex and gender differences in all stages of one's illness trajectory, but also includes the development of sex and gender-specific interventions to mitigate sex inequities in health. Sex and gender sensitivity should therefore be incorporated in health research, as sex and gender sensitive research paves the way for SGSM. This could be achieved via complementary bottom-up and top-down approaches. The former could include fostering interdisciplinary collaboration and research networks in which best practices and knowledge regarding sex and gender sensitive research can be shared among peers. The latter may include requesting obligatory sex and/or gender-stratified analyses in the policies of organizations allocating resources and editorial policies of scientific journals. Additionally, FAIR data principles (i.e., Findable, Accessible, Interoperable and Reusable data) should dictate the necessity of well-operationalized and well-defined sex and/or gender variables.

These are solely first steps in the process of a broad implementation of sex and gender sensitivity in research. After all, sex and gender sensitive research should provide (clinically) actionable results and recommendations that are tailored to the patient. For SGSM to have an effective impact on clinical practice, it is pivotal that one realizes that sex and gender are part of real-world situations, which are complex, dynamic and constantly changing.[60] Therefore, sex and gender-sensitivity should not only be integrated in epidemiological research that gathers knowledge on a population level, but also in research that focusses on gathering knowledge at the individual level.[61] Qualitative methods such as ethnography, photo-voice or focus groups may aid in gaining an in-depth understanding of the intricate interplay between sex, gender, and health in the context of other relevant factors.

However, SGSM is limited in achieving personalized and patient-centric medicine if sex and/or gender are included in research as binary or categories.[60] As aforementioned, people with intersex variations cannot be categorized into male or female, but rather form a bridge between or develop beyond these categories. The categorization of people with intersex variations into male and female sex in healthcare is problematic, as this categorization may be accompanied by the denial of their lived experiences, social stigmatization, and potential secrecy regarding their bodies.[62,63]

Therefore, it is required for SGSM to move beyond categorization of sex and/or gender.[63] But even when we consider sex and/or gender as more diverse than mere binary categories, this still disregards other health-related factors. In the end, sex and gender are only two of the factors that are part of an intricate interplay of multiple social identities. Sexual orientation, cultural and religious backgrounds, health literacy, SES, race and many more dimensions of one's identity synergistically interact and determine an

individual's health.[61] These social identities are shaped by macro-level processes, such as legislation, that affect health as well.[62] Accounting for these factors could result in increased external validity of epidemiological studies, and thus benefit the extent to which studies' results can be applied to populations beyond the study population. These intersectional approaches should be considered in all stages of epidemiological research. This includes intersectionally-informed research questions that consider whether the research is situated in certain social contexts, but also intersectionally-informed study designs that consider whether data collection or analyses categorize participants in certain ways, resulting in more comprehensive research results.[64] In data analysis, the synergistic interactions between social identities should be considered and subsequently be interpreted in light of the social, political, economic, and cultural contexts.[64]

The incorporation of intersectional approaches in epidemiology may overcome the idea of groups being homogeneous, but within epidemiology statistical models remain a simplification of real-world complexities. Therefore, more detailed, personalized, and clinically-actionable results may be obtained by applying different qualitative or participatory methodologies in addition to epidemiological methods.[61,62] An integration of the knowledge derived from these mixed-method results will allow for a more holistic understanding of what determines health and disease and more detailed guidance for policy development. To make full use of the results from these studies, an interdisciplinary research team is highly recommended.

Recommendations for policy, clinical practice and future research

Multiple recommendations for policy and clinical practice have been provided throughout this thesis and discussion. Some pivotal recommendations are explicated below, as these considerations may aid the development of more equitable and inclusive research and healthcare practices.

Policy

Sex and gender were found to be important factors at the critical junctions throughout the illness trajectory of common somatic symptoms sex. Therefore, sex and/or gender considerations should be incorporated in healthcare policies that recognize the importance and influence of these concepts on health. This may involve top-down promotion of sex and gender sensitive research by allocating resources towards research that specifically focusses on sex and/or gender differences in health and healthcare in the form of grants or fellowships. Funding agencies currently frequently

require a consideration of sex and gender in health-related research proposals, and some provide training courses on sex and gender sensitivity in research to ensure that researchers have a solid understanding of the nuances of sex and gender. Funding agencies and editorial boards could also take on a frontrunner role in sex and gender-sensitivity and require a strict adherence to the Sex And Gender Equity in Research (SAGER) guidelines within the funded and published projects. These guidelines describe a systematic method of reporting sex and gender differences in research across disciplines.[65] Adherence to uniform guidelines results in less heterogeneity in the operationalization of sex and gender, which allows for systematic reviews and meta-analyses focused on sex and/or gender differences in health. Ethical boards may also be required to explicitly consider sex and gender in their review processes, to ensure the principles of inclusivity and non-discrimination are adhered to.

Clinical practice

It is difficult to formulate recommendations aimed to change the standard of care for common somatic symptoms based on the results in this thesis, as any changes in these standards should be based on research assessing the quality of care. The research in this thesis showed the presence of sex and/or gender differences in the illness trajectories of people with common somatic symptoms, but did not clearly identify care practices that result in an inferior quality of care for women compared to men, or vice versa. Nevertheless, by being aware of sex and/or gender differences in illness trajectories of common somatic symptoms, primary care providers allow for more inclusive and patient-centric healthcare. Primary care providers could also explore the sex and gender differences in their own medical actions, such as the provision of diagnostic interventions and the effectiveness hereof, and in their communication.[66] As sex differences in medical actions may be due to gender stereotypes, primary care providers are encouraged to identify their own gender stereotypes and discuss these within the medical community in order to prevent misconceptions and provide gender-sensitive care that is tailored to the patient.

It is important for primary care providers to realize that sex and gender differences in illness trajectories for common somatic symptoms are not necessarily an indicator of a decreased quality of care. The identified sex and gender differences in this thesis may actually be indicative of patient-centric care in primary care. Based on their clinical experience, primary care providers may consider different diagnostic interventions suitable for men and women. However, primary care providers always consider the context in which patients present themselves, including the patient's (medical) history, family history and SES. A patient's sex and gender are part of this context, but may be easily overlooked.

Future research

Throughout this thesis we have provided multiple recommendations for future research. Most notable is the recommendation for research to adequately and consistently include sex and gender measures in large-scale epidemiological cohort studies. The inclusion of adequate gender measures disregards the need for gender indices and allows for more comprehensive research to assess its associations with illness trajectories. Including gender measures in longitudinal studies with multiple assessments over time would also allow for capturing the dynamics of gender over time. Currently, meta-analytic epidemiological evidence of the associations between gender and somatic symptoms is scarce, which might be partly due to the large heterogeneity in the operationalization of gender in epidemiological studies. Therefore, it should also be assessed whether a standardized operationalization of all dimensions of gender is desired in epidemiological studies, despite the concept's sensitivity to time, culture and place.

Second, we should return to the main aim of this thesis: to gain in-depth insights into whether and how sex and gender are associated with the illness trajectories of common somatic symptoms. While the majority of studies included in this thesis employ a quantitative approach to explore these differences, there would be value in incorporating qualitative and participatory approaches as well. Epidemiological, quantitative approaches allow for unravelling patterns of health that are applicable to large populations. These patterns are not necessarily applicable to the individual patient. However, qualitative approaches allow for obtaining detailed insights into the intricacies and complexities that shape individuals' illness trajectories in relation to sex and gender. In participatory action research, participants and researchers collaborate on an equal basis throughout the whole of the research process and the experiential knowledge of those experiencing inequities is considered pivotal.[67] The prioritization of practical and experiential knowledge of participants in participatory research makes this type of research specifically suitable for identification of sex inequities in illness trajectories that have a high impact on the individual, but may remain undetected by epidemiological research.[56,62] Participatory research produces a different type of knowledge than quantitative research, which could be used to identify fruitful points of intervention for policies aiming to reduce sex inequities and to evaluate such policies. In other words, qualitative and participatory research approaches might provide a deeper level of understanding of the lived experiences of people with common somatic symptoms and how they perceive the influences of their sex and/or gender on their health. In order to provide patient-centric healthcare, we need to fully understand how sex, gender, and other health-related factors interact and shape the experience and meaning of somatic symptoms in patients' lives. Therefore, qualitative research that facilitates the collection of in-depth and thick data should be initiated.

A final note

We end this thesis by reflecting on sex and gender-sensitivity in medicine. SGSM is a dynamic concept. While there is debate about whether early medical practices were truly sensitive to sex and gender differences between patients, one could argue that primitive forms of SGSM included medicine solely for the female body, including Hippocratic gynecology assuming mobility of the uterus. Over time, SGSM developed into a feminist movement that focused on improving women's wellbeing, their health, and healthcare. Currently, it is evolved into a global cooperative community of healthcare professionals, researchers and activists all united to pursue optimal and inclusive healthcare for all genders and sexes. We find that at every twist and turn in people's illness trajectory sex and gender play a role. Ultimately, only by collaborative efforts between researchers and those experiencing the influences of sex and gender on their illness trajectories the most comprehensive roadmap for illness trajectories can be sketched.

References

1. Ballering AV, Ori APS, Rosmalen JGM. The association of sex, age and FKBP5 genotype with common somatic symptoms: A replication study in the lifelines cohort study. *Journal of psychosomatic research*. 2021;147:110510. doi: 10.1016/j.jpsychores.2021.110510.
2. Hogendoorn E, Ballering AV, van Dijk MWG, Rosmalen JGM, Burke SM. Discordance between adolescents and parents in functional somatic symptom reports: Sex differences and future symptom prevalence. *Journal of Youth and Adolescence*. 2023. doi: 10.1007/s10964-023-01810-w.
3. Ballering AV, Bonvanie IJ, Olde Hartman TC, Monden R, Rosmalen JGM. Gender and sex independently associate with common somatic symptoms and lifetime prevalence of chronic disease. *Social Science & Medicine*. 2020;253:112968. doi: 10.1016/j.socscimed.2020.112968.
4. Ballering AV, Wardenaar KJ, olde Hartman TC, Rosmalen JGM. Female sex and femininity independently associate with common somatic symptom trajectories. *Psychological Medicine*. 2022;52(11):2144-2154. doi: 10.1017/S0033291720004043.
5. Ballering AV, van Zon SKR, olde Hartman TC, Rosmalen JGM. Persistence of somatic symptoms after COVID-19 in the netherlands: An observational cohort study. *The Lancet*. 2022;400(10350):452-461. doi: 10.1016/S0140-6736(22)01214-4.
6. Ballering AV, Olde Hartman TC, Verheij R, Rosmalen JGM. Sex and gender differences in primary care help-seeking for common somatic symptoms: A longitudinal study. *Scandinavian Journal of Primary Health Care*. 2023;41(2):132-139. doi: 10.1080/02813432.2023.2191653.
7. Ballering AV, Muijres D, Uijen AA, Rosmalen JGM, olde Hartman TC. Sex differences in the trajectories to diagnosis of patients presenting with common somatic symptoms in primary care: An observational cohort study. *Journal of Psychosomatic Research*. 2021;149:110589. doi: 10.1016/j.jpsychores.2021.110589.
8. Groeneveld JM, Ballering AV, van Boven K, Akkermans RP, Olde Hartman TC, Uijen AA. Sex differences in incidence of respiratory symptoms and management by general practitioners. *Family practice*. 2020;37(5):631-636. doi: 10.1093/fampra/cmaa040.
9. Ballering A, Rosmalen J, Olde Hartman T. Differences between women and men are present in the rate of diagnosed diseases after a diagnostic intervention is conducted in primary care. *Journal of the American Board of Family Medicine*. 2022;35(1):73-84. doi: 10.3122/jabfm.2022.01.210289.
10. Ballering AV, Oertelt-Prigione S, LCRI, Olde Hartman TC, Rosmalen JGM. Sex and gender-related differences in COVID-19 diagnoses and SARS-CoV-2 testing practices during the first wave of the pandemic: The dutch lifelines COVID-19 cohort study. *Journal of Women's Health (Larchmt.)*. 2021;30(12):1686-1692. doi: 10.1089/jwh.2021.0226.
11. Ballering AV, Oertelt-Prigione S, Olde Hartman TC, Rosmalen JGM. Response to rossato et al. *Journal of Women's Health (Larchmt.)*. 2022;31(6):896-898. doi: 10.1089/jwh.2022.0157.
12. Ballering AV, Burke SM, Maeckelberghe ELM, Rosmalen JGM. How to ensure inclusivity in large-scale general population cohort studies? lessons learned with regard to including and assessing sex, gender, and sexual orientation. *Archives of Sexual Behavior*. 2023(52):2163–2172. doi: 10.1007/s10508-023-02600-y.
13. Ballering AV, Plug I, Lagro-Janssen ALM, Das E, Rosmalen JGM. Building a bridge between multidisciplinary insights and practice: The development of an e-learning for internal residents about sex, gender and persistent somatic symptoms. *Journal of Psychosomatic Research*. 2022;155:110739. doi: 10.1016/j.jpsychores.2022.110739.

14. Klijs B, Scholtens S, Mandemakers JJ, Snieder H, Stolk RP, Smidt N. Representativeness of the LifeLines cohort study. *PLoS ONE*. 2015;10(9):e0137203. doi: 10.1371/journal.pone.0137203.
15. Oldehinkel AJ, Rosmalen JG, Buitelaar JK, et al. Cohort profile update: The tracking adolescents' individual lives survey (TRAILS). *International Journal of Epidemiology*. 2015;44(1):76-76n. doi: 10.1093/ije/dyu225.
16. Luijks H, van Boven K, olde Hartman T, Uijen A, van Weel C, Schers H. Purposeful incorporation of patient narratives in the medical record in the netherlands. *Journal of the American Board of Family Medicine*. 2021;34(4):709-723. doi: 10.3122/jabfm.2021.04.200609.
17. Zijlema WL, Stolk RP, Löwe B, Rief W, White PD, Rosmalen JGM. How to assess common somatic symptoms in large-scale studies: A systematic review of questionnaires. *Journal of Psychosomatic Research*. 2013;74(6):459-68. doi: 10.1016/j.jpsychores.2013.03.093.
18. Soler JK, Okkes I. Reasons for encounter and symptom diagnoses: A superior description of patients' problems in contrast to medically unexplained symptoms (MUS). *Family practice*. 2012;29(3):272-282. doi: 10.1093/fampra/cmr101.
19. Erik Børve Rasmussen. Making and managing medical anomalies: Exploring the classification of 'medically unexplained symptoms'. *Social Studies of Science*. 2020;50(6):901-931. doi: 10.1177/0306312720940405.
20. Miron-Shatz T, Ormianer M, Rabinowitz J, Hanoch Y, Tsafrir A. Physician experience is associated with greater underestimation of patient pain. *Patient Education and Counseling*. 2020;103(2):405-409. doi: 10.1016/j.pec.2019.08.040.
21. Samulowitz A, Gremyr I, Eriksson E, Hensing G. "Brave men" and "Emotional women": A theory-guided literature review on gender bias in health care and gendered norms towards patients with chronic pain. *Pain Research and Management*. 2018;2018:6358624. doi: 10.1155/2018/6358624.
22. Pedhazur EJ, Tetenbaum TJ. Bem sex role inventory: A theoretical and methodological critique. *Journal of Personality and Social Psychology*. 1979;37(6):996-1016. doi: 10.1037/0022-3514.37.6.996.
23. Hoffman RM, Borders LD. Twenty-five years after the bem sex-role inventory: A reassessment and new issues regarding classification variability. *Measurement and Evaluation in Counseling and Development*. 2001;34(1):39-55. doi: 10.1080/07481756.2001.12069021.
24. Horstmann S, Schmechel C, Palm K, Oertelt-Prigione S, Bolte G. The operationalisation of sex and gender in quantitative Healthâ€"Related research: A scoping review. *International Journal of Environmental Research and Public Health*. 2022;19(12). doi: 10.3390/ijerph19127493.
25. Bem SL. The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*. 1974;42(2):155-162. doi: 10.1037/h0036215.
26. Donnelly K, Twenge JM. Masculine and feminine traits on the bem sex-role inventory, 1993–2012: A cross-temporal meta-analysis. *Sex Roles*. 2017;76(9):556-565.
27. Pelletier R, Ditto BF, Pilote L. A composite measure of gender and its association with risk factors in patients with premature acute coronary syndrome. *Psychosomatic medicine*. 2015;77(5):517-526. doi: 10.1097/PSY.0000000000000186.
28. Vader SS, Lewis SM, Verdonk P, Verschuren WMM, Picavet HSJ. Masculine gender affects sex differences in the prevalence of chronic health problems - the doetinchem cohort study. *Prev Med Rep*. 2023;33:102202. doi: 10.1016/j.pmedr.2023.102202.
29. Nauman AT, Behlouli H, Alexander N, et al. Gender score development in the berlin aging study II: A retrospective approach. *Biology of Sex Differences*. 2021;12(1):15. doi: 10.1186/s13293-020-00351-2.

30. Levinsson A, de Denus S, Sandoval J, et al. Construction of a femininity score in the UK biobank and its association with angina diagnosis prior to myocardial infarction. *Scientific Reports*. 2022;12(1):1780. doi: 10.1038/s41598-022-05713-x.
31. Lacasse A, Pagé MG, Choinière M, et al. Conducting gender-based analysis of existing databases when self-reported gender data are unavailable: The GENDER index in a working population. *Canadian Journal of Public Health*. 2020;111(2):155-168. doi: 10.17269/s41997-019-00277-2.
32. Smith PM, Koehoorn M. Measuring gender when you don't have a gender measure: Constructing a gender index using survey data. *International Journal for Equity in Health*. 2016;74(6):459-468. doi: 10.1186/s12939-016-0370-4.
33. Al-Hassany L, Verdonk P, Maassen Van den Brink A. Studying sex and gender in neurological disease. *Lancet Neurology*. 2023;22(1):32-33. doi: 10.1016/S1474-4422(22)00493-8.
34. Traag T. Opleiding en werk: Twee generaties vrouwen vergeleken over deze public. https://www.cbs.nl/nl-nl/longread/statistische-trends/2020/opleiding-en-werk-twee-generaties-vrouwen-vergeleken. Updated 2020. Accessed 01 August, 2023.
35. Lippa R, Connelly S. Gender diagnosticity: A new bayesian approach to gender-related individual differences. *Journal of Personality and Social Psychology*. 1990;59(5):1051-1065. doi: 10.1037/0022-3514.59.5.1051.
36. Weber AM, Gupta R, Abdalla S, Cislaghi B, Meausoone V, Darmstadt GL. Gender-related data missingness, imbalance and bias in global health surveys. *BMJ Glob Health*. 2021;6(11):e007405. doi: 10.1136/bmjgh-2021-007405.
37. Cislaghi B, Weber AM, Shakya HB, et al. Innovative methods to analyse the impact of gender norms on adolescent health using global health survey data. *Social Science & Medicine*. 2022;293:114652. doi: 10.1016/j.socscimed.2021.114652.
38. World values survey: Round Seven–Country-pooled datafile version 5.0. In: Haerpfer C, Inglehart R, Moreno A, et al, eds. Madrid, Spain & Vienna, Austria: JD Systems Institute & WVSA Secretariat; 2022. 10.14281/18241.20.
39. Bolijn R, Anton EK, Appelman Y, et al. Prospective analysis of gender-related characteristics in relation to cardiovascular disease. *Heart*. 2022;108(13):1030-1038. doi: 10.1136/heartjnl-2021-320414.
40. Bolijn R, Perini W, Tan HL, Galenkamp H, Kunst AE, van Valkengoed IGM. Gender-related characteristics and disparities in estimated cardiovascular disease risk in a multi-ethnic general population: The HELIUS study. *International Journal of Cardiology*. 2021;327:193-200. doi: 10.1016/j.ijcard.2020.11.041.
41. Voices for Women. Meldpunt - voices for women. https://voicesforwomen.nl/meldpunt/. Updated 2022. Accessed 22 June, 2023.
42. Chaabouni A, Houwen J, Walraven I, et al. Patients' characteristics and general practitioners' management of patients with symptom diagnoses: A retrospective cohort study. *Journal of the American Board of Family Medicine*. 2023(36):477-492.
43. Lutfey KE, Link CL, Marceau LD, et al. Diagnostic certainty as a source of medical practice variation in coronary heart disease: Results from a cross-national experiment of clinical decision making. *Medical Decision Making*. 2009;29(5):606-618. doi: 10.1177/0272989X09331811.
44. van der Weijden T, van Velsen M, Dinant G, van Hasselt CM, Grol R. Unexplained complaints in general practice: Prevalence, patients' expectations, and professionals' test-ordering behavior. *Medical Decision Making*. 2003;23(3):226-231.
45. Michiels-Corsten M, Donner-Banzhoff N. Beyond accuracy: Hidden motives in diagnostic testing. *Family Practice*. 2018;35(2):222-227.

46. Plug I, Stommel W, Lucassen P, Olde Hartman T, Van Dulmen S, Das E. Do women and men use language differently in spoken face-to-face interaction? A scoping review. *Review of Communication Research*. 2021;9:43-79. doi: 10.12840/ISSN.2255-4165.026.
47. Hall JA, Roter DL. Do patients talk differently to male and female physicians?: A meta-analytic review. *Patient Education and Counseling*. 2002;48(3):217-224. doi: 10.1016/S0738-3991(02)00174-X.
48. Sandhu H, Adams A, Singleton L, Clark-Carter D, Kidd J. The impact of gender dyads on doctor–patient communication: A systematic review. *Patient Education and Counseling*. 2009;76(3):348-355. doi: 10.1016/j.pec.2009.07.010.
49. Plug I, van Dulmen S, Stommel W, Olde Hartman TC, Das E. Physicians' and patients' interruptions in clinical practice: A quantitative analysis. *Annals of Family Medicine*. 2022;20(5):423-429. doi: 10.1370/afm.2846.
50. Lindvall-Östling M, Deutschmann M, Steinvall A. An exploratory study on linguistic gender stereotypes and their effects on perception. *Open Linguistics*. 2020;6(1):567-583. doi: 10.1515/opli-2020-0033.
51. Claréus B, Renström EA. Physicians' gender bias in the diagnostic assessment of medically unexplained symptoms and its effect on patient–physician relations. *Scandinavian Journal of Psychology*. 2019;60(4):338-347. doi: 10.1111/sjop.12545.
52. Plug I. *A linguistic view on gender in medical interaction.* [PhD]. Radboud University; 2022.
53. Atasoy S, Hausteiner-Wiehle C, Sattel H, et al. Gender specific somatic symptom burden and mortality risk in the general population. *Scientific reports*. 2022;12(1):15049. doi: 10.1038/s41598-022-18814-4.
54. MacLean A, Hunt K, Smith S, Wyke S. Does gender matter? an analysis of men's and women's accounts of responding to symptoms of lung cancer. *Social Science & Medicine*. 2017;191:134-142. doi: 10.1016/j.socscimed.2017.09.015.
55. Mauvais-Jarvis F, Bairey Merz N, Barnes PJ, et al. Sex and gender: Modifiers of health, disease, and medicine. *The Lancet*. 2020;396(10250):565-582. doi: 10.1016/S0140-6736(20)31561-0.
56. Dijkstra I, Horstman K. 'Known to be unhealthy': Exploring how social epidemiological research constructs the category of low socioeconomic status. *Social Science & Medicine*. 2021;285:114263. doi: 10.1016/j.socscimed.2021.114263.
57. Cleghorn E. *Unwell women - A journey through medicine and myth in a man-made world.* 1st ed. London: Weidenfeld & Nicolson; 2021.
58. Becher E, Oertelt-Prigione S. History and development of sex- and gender sensitive medicine (SGSM). *International Review of Neurobiology*. 2022;164:1-25. doi: 10.1016/bs.irn.2022.06.008.
59. Nowatzki N, Grant KR. Sex is not enough: The need for gender-based analysis in health research. *Health Care for Women International*. 2011;32(4):263-277. doi: 10.1080/07399332.2010.519838.
60. Mena E, Bolte G, on behalf of the ADVANCE GENDER Study Group. Intersectionality-based quantitative health research and sex/gender sensitivity: A scoping review. *International Journal for Equity in Health*. 2019;18(1):199. doi: 10.1186/s12939-019-1098-8.
61. Göttgens I, Oertelt-Prigione S. Moving beyond gender identity: The need for contextualization in gender-sensitive medical research. *The Lancet Regional Health – Europe*. 2023;24. doi: 10.1016/j.lanepe.2022.100548.
62. Verdonk P, Muntinga M, Leyerzapf H, Abma T. From gender sensitivity to an intersectionality and participatory approach in health research and public policy in the netherlands. In: Hankivsky O, Jordan-Zachery JS, eds. *The palgrave handbook of intersectionality in public policy.* London: Springer International Publishing; 2019:413-432. 10.1007/978-3-319-98473-5_18.

63. Davy Z, Siriwardena AN. To be or not to be LGBT in primary health care: Health care for lesbian, gay, bisexual, and transgender people. *British Journal of General Practice*. 2012;62(602):491-492. doi: 10.3399/bjgp12X654731.
64. Rouhani S. Intersectionality-informed quantitative research: A primer. *Institute for intersectionality Research and policy, SFU*. 2014:1-14.
65. Heidari S, Babor TF, De Castro P, Tort S, Curno M. Sex and gender equity in research: Rationale for the SAGER guidelines and recommended use. *Research Integrity and Peer Review*. 2016;1:2. doi: 10.1186/s41073-016-0007-6.
66. Stommel W, Plug I, olde Hartman TC, Lucassen PLBJ, van Dulmen S, Das E. Gender stereotyping in medical interaction: A membership categorization analysis. *Patient Education and Counseling*. 2022;105(11):3242-3248. doi: 10.1016/j.pec.2022.07.018.
67. Cornish F, Breton N, Moreno-Tabarez U, et al. Participatory action research. *Nature Reviews Methods Primers*. 2023;3(1):34. doi: 10.1038/s43586-023-00214-1.

APPENDICES

SUMMARY

Summary

Every once in a while, everyone experiences a symptom such as headache, nausea and tiredness. Usually these symptoms disappear spontaneously, but symptoms may persist. Illness trajectories for these common somatic symptoms may be initiated when a person notices a bodily sensation that is interpreted to be "wrong". An illness trajectory only ends if the idea of a sensation being wrong and concomitant care that are provided for the symptom are deemed to be finished. During an illness trajectory, which may last from a mere few hours to many years, multiple critical junctions occur. Critical junctions are turning points involving an experience or event, or interdependent sequence of events, that have a potentially far-fetching impact on the patient's health and healthcare experience. These critical junctions include among others the interpretation of the severity of a symptom, seeking healthcare for symptoms, and being provided with a diagnosis for the symptoms.

Previous research has shown that throughout illness trajectories of a variety of health conditions sex and gender differences are present at many critical junctions. Despite the body of evidence that exists on sex and gender differences in health in general, knowledge on these differences in the illness trajectories of common somatic symptoms is lacking. Therefore, this thesis aims to gain in-depth insights into whether and how sex and gender associate with the illness trajectories of common somatic symptoms, while taking into account the multifaceted nature of sex and gender.

1. Etiology of common somatic symptoms
This thesis starts by exploring the first critical junction in an illness trajectory, namely the interpretation of a bodily sensation as a symptom. To this end, we study the biological and psychosocial factors in which the development of common somatic symptoms are rooted. Women and men differ herein. In *Chapter 2* we show that sex differences occur in the biological factors contributing to common somatic symptoms. We aimed to replicate a small-scale study that reported a significant association between the rs9470080 genotype and common somatic symptom levels in women, but not in men. We could not replicate these results in the large-scale general population cohort study Lifelines, as the rs9470080 genotype did not associate with common somatic symptom levels in male or female participants. We also quantified the genetic contribution to phenotypic variation in common somatic symptom levels and found that this was higher in male than in female participants.

Not only genetic factors, but also childhood experiences affect one's proneness to common somatic symptoms. We continued the thesis by assessing whether parental assessment of adolescent's somatic symptom burden associates with symptom experience during adulthood. In *Chapter 3* we used data derived from the prospective Tracking Adolescents' Individual Lives (TRAILS) study. We found that parents of adolescent boys reported more symptoms than the boys themselves, while parents of adolescent girls reported fewer symptoms than the girls themselves. The level of parental-reported somatic symptoms during adolescence associated significantly with somatic symptom experience in adulthood. The strength of this association between parental-reported symptoms and symptoms in adulthood, did not differ over levels of adolescents' self-reported somatic symptoms or their sex.

2. Prevalence and persistence of common somatic symptoms
In adulthood, biological and psychosocial factors remain important in common somatic symptoms, as we show in the second part of this thesis. *Chapter 4* examined the cross-sectional, independent associations between sex, gender, and the prevalence of common somatic symptoms. This study was based on Lifelines data. It describes a data-driven method to develop a gender measure (i.e., the gender index) in a cohort that lacked information on participants' gender. We found that both female sex and feminine gender associated with common somatic symptoms and chronic disease. Feminine gender associated more strongly with common somatic symptoms in male participants than in female participants.

Subsequently we assessed the independent associations between sex, gender and common somatic symptoms trajectories in a longitudinal setting in *Chapter 5* by using Lifelines data. We identified five different linear trajectories of common somatic symptom severity over time in a data-driven manner. These varied from a stable and low severity trajectory in the vast majority of the general population, to a trajectory with increasing symptom severity in relatively few people. When comparing these groups, we found that female sex positively associated with a higher symptom severity, whereas feminine gender negatively associated with this.

In *Chapter 6* we described the trajectory of 23 somatic symptoms surrounding a COVID-19 diagnosis. Due to the unique nature of the Lifelines COVID-19 Cohort Study we could account for individuals' symptom levels before their COVID-19 diagnosis, as well as symptom levels among an uninfected control population. We identified three types of symptoms: (1) acute symptoms in which symptom severity was worsened in the days surrounding the COVID-19 diagnosis, but returned to pre-COVID-19 levels within 90 days; (2) core symptoms in which symptom severity persisted and did not

return to pre-COVID-19 levels within 90 days and; (3) other symptoms, for which no distinctive pattern was observed. Visual inspection of the data showed that during the acute phase of COVID-19, women reported higher severity for the majority of acute symptoms and core symptoms. We considered both female sex and feminine gender to be important in this female preponderance in symptom experience.

3. Primary care help-seeking for common somatic symptoms
Another critical junction in the illness trajectory for common somatic symptoms is the decision to seek help from a healthcare professional for one's symptoms. We studied potential sex and gender differences in this critical junction in the third section of this thesis (*Chapter 7*). To this end we linked data from Lifelines with the Nivel Primary Care Database. We show that female sex rather than feminine gender, operationalized by the gender index, was associated with consulting the general practitioner (GP) for common somatic symptoms. People's number of paid working days was negatively associated with primary care help-seeking for common somatic symptoms. Our results suggest that factors related to the frequency of help-seeking are rooted in biology, or in gendered components that are not sufficiently reflected by the composite gender index.

4. Diagnostics in primary care for common somatic symptoms
After the decision to seek help, another critical junction takes place in the consulting room of the GP, namely whether the patient is provided with diagnostic interventions. In the fourth section of this thesis we report on studies that are based on data from the practice-based research network FaMe-Net (*Chapter 8* to *10*). We show that male patients who visited their GP with cough, dyspnea or other common somatic symptoms, were provided with more diagnostic interventions, such as a physical examination, diagnostic imaging, and a referral to a specialist than female patients. In contrast, female patients with common somatic symptoms were more frequently provided with laboratory diagnostics than male patients. These sex differences in diagnostic interventions contribute to the 6%-lower odds of women's symptoms being diagnosed with a disease compared to men. We argue that many factors, including biological processes, gendered stereotypes, communication between the GP and patients, and GP's diagnostic uncertainty, interact and partially explain the sex differences in provided interventions.

In *Chapter 10* we show that irrespective of their frequency of being provided, diagnostic interventions are less effective in contributing to a disease diagnosis in female patients than in male patients. This may also partly explain why GPs provide fewer diagnostic interventions to female patients than to male patients.

Similarly, in the COVID-19 pandemic sex and gender differences in diagnostics were observed as well (*Chapter 11*). Although we found no sex difference in SARS-CoV-2 testing practices and COVID-19 diagnosis in the general population, within the population of healthcare workers, women were less frequently tested and diagnosed than men. This male preponderance in testing and diagnosis could be explained by the more pronounced acute COVID-19 symptoms and worse prognosis, in men compared to women (*Chapter 11A*).

5. Adequate conceptual assessment and implementation of knowledge
Finally, when reflecting on the studies included in this thesis, we identified pitfalls when conducting research into health-related sex and gender differences, which we describe in *Chapter 12*. We also provide strategies and recommendations for inclusive assessment of sex and gender measures in large-scale cohort studies.

Then, in *Chapter 13* we describe the development of an e-learning course that synthesizes the research results of our interdisciplinary research project. Via this e-learning course we aim to encourage internists in training as well as their supervisors to be aware of sex and gender differences in illness trajectories of people with common somatic symptoms.

Last, in the general discussion of this thesis we summarize our main findings and explicate three themes that arose from the research findings. First, we discuss the assessment of gender in epidemiological studies, and discuss the advantages and disadvantages of gender indices in-depth. Second, we reflect on gender inequities in the diagnoses for common somatic symptoms by explicating and deconstructing underlying assumptions. Third, we discuss the future of SGSM and how intersectionality may play a role to achieve sex and gender sensitivity in research and medicine. This thesis ends with recommendations for policy, clinical practice and future research.

SAMENVATTING

Samenvatting

Iedereen ervaart zo nu en dan lichamelijke klachten, zoals hoofdpijn, misselijkheid en vermoeidheid. Meestal gaan deze veelvoorkomende lichamelijke klachten vanzelf over, maar soms houden de klachten aan. Deze klachten kunnen het begin zijn van een ziekte traject. Een ziekte traject begint wanneer iemand een lichamelijke sensatie niet goed kan duiden en eindigt op het moment dat de klacht weg is en de bijbehorende zorg gestopt is. Tijdens zo'n ziekte traject, dat enkele uren tot jaren kan duren, zijn er meerdere belangrijke momenten die de gezondheid en zorg van mensen kunnen beïnvloeden. Deze momenten noemen we kritieke kantelpunten, zoals het interpreteren van de ernst van een klacht, het zoeken van medische hulp, en het al dan niet krijgen van een diagnose.

Eerdere onderzoeken hebben laten zien dat er verschillen zijn tussen mannen en vrouwen op deze kritieke kantelpunten in ziekte trajecten. Ondanks dat er al veel kennis is over gender- en geslachtsverschillen in gezondheid, weten we nog weinig over gender- en gezondheidsverschillen in de ziekte trajecten van veelvoorkomende lichamelijke klachten. Het doel van dit proefschrift is om inzicht te krijgen in de mogelijke geslachts- en genderverschillen in de ziekte trajecten van veelvoorkomende lichamelijke klachten.

1. De etiologie van veelvoorkomende lichamelijke klachten
Dit proefschrift begint met het verkennen van het eerste kritieke kantelpunt in een ziekte traject, namelijk de interpretatie van een lichamelijke sensatie als symptoom. We onderzoeken zowel de biologische als psychosociale factoren die van belang zijn in het ontstaan van lichamelijke klachten. In *Hoofdstuk 2* laten we zien dat er geslachtsverschillen zijn in de genetica die ten grondslag ligt aan lichamelijke klachten. Hoewel een kleinschalige studie een significante associatie vond tussen het rs9470080 genotype en het vóórkomen van lichamelijke klachten bij vrouwen, maar niet bij mannen, konden wij deze resultaten niet repliceren in het grote, algemene bevolkingscohort Lifelines. Zowel bij mannen als bij vrouwen vonden wij geen associatie tussen het rs9470080 genotype en lichamelijke klachten. We kwantificeerden ook de genetische bijdrage aan fenotypische variatie in veelvoorkomende lichamelijke klachten en vonden dat deze bijdrage significant hoger was bij mannelijke deelnemers dan bij vrouwelijke deelnemers.

Niet alleen genetische factoren, maar ook ervaringen in de kindertijd dragen bij aan het ervaren van lichamelijke klachten. We onderzochten of de ouderrapportage van lichamelijke klachten bij adolescenten samenhangt met symptoomervaring op

volwassen leeftijd. Voor *Hoofdstuk 3* gebruikten we gegevens uit de prospectieve Tracking Adolescents' Individual Lives (TRAILS) studie. We laten zien dat ouders van adolescente jongens meer klachten rapporteerden dan de jongens zelf, terwijl ouders van adolescente meisjes minder klachten rapporteerden dan de meisjes zelf. De ernst van door ouders gerapporteerde klachten tijdens de adolescentie hing samen met het ervaren van lichamelijke klachten op latere leeftijd, maar de sterkte van deze samenhang verschilde niet tussen zelf gerapporteerde ernst van klachten van adolescenten of hun geslacht.

2. Prevalentie en persisteren van veelvoorkomende lichamelijke klachten
Ook op volwassen leeftijd zijn biologische en psychosociale factoren belangrijk bij veelvoorkomende lichamelijke klachten. Dit laten we zien in het tweede deel van dit proefschrift. *Hoofdstuk 4* onderzocht cross-sectioneel de samenhang tussen geslacht, gender en het vóórkomen van veelvoorkomende lichamelijke klachten. Dit onderzoek is gebaseerd op Lifelines data. Het beschrijft een data gedreven methode om een gendermaat te ontwikkelen. Dit deden we in een cohort waarin geen directe informatie over gender beschikbaar is. We vonden dat zowel vrouwelijk geslacht als vrouwelijk gender samenhing met het vóórkomen van veelvoorkomende lichamelijke klachten en chronische ziekten. Vrouwelijk gender hing sterker samen met veelvoorkomende lichamelijke klachten bij mannen dan bij vrouwen.

Vervolgens onderzochten we in *Hoofdstuk 5* de samenhang tussen geslacht, gender en het beloop van veelvoorkomende lichamelijke klachten in een longitudinale setting. We identificeerden, via een data gedreven methode, vijf verschillende lineaire trajecten die de ernst van veelvoorkomende lichamelijke klachten over tijd beschrijven. De meeste mensen ervaarden weinig klachten over tijd, maar een kleine groep had sterk toenemende klachten in de loop van de tijd. Wanneer we deze groepen vergelijken, zien we dat vrouwelijk geslacht positief geassocieerd is met een ernstiger beloop van klachten, terwijl vrouwelijk gender negatief geassocieerd is met een ernstiger beloop.

In *Hoofdstuk 6* beschreven we het beloop van 23 lichamelijke klachten rondom een COVID-19 diagnose. Vanwege de unieke aard van de Lifelines COVID-19 Cohort Studie konden we rekening houden met het niveau van klachten van individuen vóór de COVID-19-diagnose, en met het niveau van klachten in een niet geïnfecteerde controlegroep. We vonden drie types klachten: (1) acute klachten waarbij de ernst van de klachten verergerde in de dagen rondom de COVID-19-diagnose, maar binnen 90 dagen terugkeerde naar het niveau van vóór COVID-19 en naar het niveau in de niet geïnfecteerde controlegroep; (2) kernklachten waarbij de ernst van de klachten

aanhield en niet binnen 90 dagen terugkeerde naar de niveaus van vóór COVID-19 of naar het niveau in de controlegroep; en (3) andere klachten, waarvoor we geen kenmerkend patroon voor COVID-19 patiënten ten opzichte van vóór COVID-19 of ten opzichte van de controlegroep vonden. Visuele inspectie van het beloop van deze lichamelijke klachten liet zien dat vrouwen ernstigere en langer aanhoudende klachten ervaarden dan mannen. We beschouwen zowel vrouwelijk geslacht en vrouwelijk gender als van belang in het ervaren van klachten rondom COVID-19.

3. Hulp zoeken bij de huisarts voor veelvoorkomende lichamelijke klachten
Een volgend kantelpunt in het ziekte traject van veelvoorkomende lichamelijke klachten is de beslissing om hulp te zoeken bij de huisarts. We onderzochten de mogelijke gender- en geslachtsverschillen in het zoeken van hulp voor lichamelijke klachten in het derde deel van dit proefschrift. We hebben Lifelines data gekoppeld aan de Nivel Primary Care Database. In *Hoofdstuk 7* laten we zien dat vrouwelijk geslacht, maar niet vrouwelijk gender, samenhangt met het zoeken van hulp voor lichamelijke klachten. Specifieke gender-gerelateerde factoren, zoals het aantal dagen dat iemand betaald werk uitvoert in de week, voorspelden het zoeken van hulp. Onze resultaten suggereren dat biologische factoren, of factoren die niet goed gereflecteerd worden door de gender index, samenhangen met het zoeken van hulp bij de huisarts voor lichamelijke klachten.

4. Diagnostiek in de eerstelijnszorg voor veelvoorkomende lichamelijke klachten
Nadat iemand de beslissing heeft genomen om hulp te zoeken voor lichamelijke klachten, volgt een volgend kantelpunt in het ziekte traject, namelijk in de spreekkamer van de huisarts, waar de patiënt al dan niet diagnostische interventies aangeboden krijgt. In het vierde deel van dit proefschrift beschrijven we onderzoeken met behulp van data uit het praktijkgerichte onderzoeksnetwerk FaMe-Net (*Hoofdstuk 8* tot en met *10*). We tonen aan dat mannelijke patiënten die zich presenteerden met een hoest, kortademigheid of andere veelvoorkomende lichamelijke klachten meer diagnostische interventies krijgen van de huisarts dan vrouwelijke patiënten. Denk hierbij aan lichamelijk onderzoek, diagnostische beeldvorming en een verwijzing naar een specialist. Vrouwelijke patiënten met lichamelijke klachten kregen daarentegen vaker laboratoriumdiagnostiek dan mannelijke patiënten. Deze geslachtsverschillen in verstrekte diagnostische interventies dragen bij aan de 6% lagere kans van vrouwen ten opzichte van mannen om gediagnosticeerd te worden met een ziekte wanneer ze met lichamelijke klachten bij de huisarts komen. Wij denken dat veel factoren, waaronder biologische processen, genderstereotypen, communicatie(stijl), en diagnostische onzekerheid, deels met deze man-vrouw verschillen te maken hebben.

In *Hoofdstuk 10* laten we zien dat diagnostische interventies minder effectief bijdragen aan een ziektediagnose bij vrouwen dan bij mannen. Ook dit kan gedeeltelijk verklaren waarom huisartsen minder diagnostische interventies bij vrouwen doen, dan bij mannen.

Ook gedurende de eerste fase van de COVID-19 pandemie zagen we man-vrouw verschillen in diagnostiek. In *Hoofdstuk 11* onderzochten we of er verschillen zijn in het testen op SARS-CoV-2 besmetting en COVID-19 diagnoses tijdens de eerste golf van de pandemie. Hoewel we geen man-vrouw verschillen vonden in het testen op SARS-CoV-2 en in COVID-19 diagnoses in de algemene bevolking, vonden we deze wel bij zorgmedewerkers. Vrouwelijke zorgmedewerkers werden minder vaak getest en gediagnosticeerd dan mannelijke zorgmedewerkers. Dit kan deels verklaard worden door de ernstigere COVID-19-symptomen die mannen ervaren in vergelijking met vrouwen (*Hoofdstuk 11A*).

5. Het adequaat uitvragen van concepten en kennisimplementatie
Als laatste reflecteren we kritisch op de onderzoeken in dit proefschrift. We merkten een aantal valkuilen op tijdens ons onderzoek naar geslachts- en genderverschillen in gezondheid. Deze valkuilen beschrijven we in *Hoofdstuk 12* en we geven aanbevelingen om deze valkuilen te omzeilen. Bovendien bevelen we concrete strategieën aan voor grootschalige cohortstudies om inclusiever te worden met betrekking tot geslacht, gender en seksuele oriëntatie.

In *Hoofdstuk 13* beschrijven we de ontwikkeling van een e-learning die de onderzoeksresultaten van ons interdisciplinaire onderzoeksproject samenbrengt. Via de e-learning willen we internisten in opleiding en hun begeleiders stimuleren om zich bewust te zijn van geslachts- en genderverschillen in ziekte trajecten van patiënten met veelvoorkomende lichamelijke klachten.

In de algemene discussie van dit proefschrift vatten we onze belangrijkste bevindingen samen en diepen we drie thema's uit die voortkwamen uit de onderzoeksresultaten. Ten eerste bespreken we het meten van gender in epidemiologische studies en de voor- en nadelen van gender indices. Ten tweede reflecteren we op genderongelijkheden bij de diagnoses voor veelvoorkomende lichamelijke klachten door onderliggende aannames over gender- en geslachtsverschillen expliciet te maken en indien nodig te weerleggen. Ten derde bespreken we de toekomst van geslacht- en gendersensitieve geneeskunde en hoe intersectionaliteit een rol kan spelen in de toekomst van geslacht en gendersensitiviteit in onderzoek en geneeskunde. We eindigen dit proefschrift met het formuleren van aanbevelingen voor beleid, de klinische praktijk en toekomstig onderzoek.

ABOUT THE AUTHOR

About the author

Aranka Vivienne Ballering was born in Alphen aan den Rijn, 1994. She finished the gymnasium at the Groene Hart Lyceum. After obtaining her bachelor's degree in Health and Life Sciences (*cum laude*) at VU University Amsterdam, Aranka continued with the research master Biomedical Sciences (*cum laude*) at the same university. She specialized in immunology and international public health. During her studies she spent five months in Varanasi (India), to conduct research assessing the stigmatization of people affected by leprosy. Thereafter, she obtained a second master's degree in Global Health from Maastricht University (*cum laude*). Her master thesis comprised research exploring the knowledge, attitudes and practices of the general public towards people living with HIV/AIDS. From 2019 to 2023, Aranka worked on a PhD project focusing on the associations between sex, gender and common somatic symptoms at the UMCG. She was supervised by Judith Rosmalen (UMCG) and Tim Olde Hartman (Radboudumc), but also collaborated closely with clinical psychologists, sociologists, pedagogical specialists, and language and communication specialists. Aranka was involved in the supervision of multiple students and presented her work at international and national conferences. During her PhD trajectory, Aranka participated in a variety of knowledge utilization activities, including the development of a game of quartet, the 3-Minute-Thesis competition, the Hoe?Zo! Show, and the Royal Netherlands Academy of Arts and Sciences' Faces of Science program. She interacted with the press via interviews for radio, television, newspapers, and podcasts. She currently works as a researcher at the Department of Psychiatry at the UMCG focusing on intergenerational transmission of symptom proneness using, among others, qualitative research methodologies. Ultimately, Aranka aims to pursue research exploring the stigmatization of people with persistent somatic symptoms.

LIST OF PUBLICATIONS

Included in this thesis

Hogendoorn, E., **Ballering, A.V.**, van Dijk, M.W.G., Rosmalen, J.G.M., Burke, S.M. (2023) Discordance between Adolescents and Parents in Functional Somatic Symptom Reports: Sex Differences and Future Symptom Prevalence. *Journal of Youth and Adolescence,* 52, 2182–2195 (Chapter 3)

Ballering, A.V., Burke, S.M., Maeckelberghe, E.L.M., Rosmalen, J.G.M. (2023) How to ensure inclusivity in large-scale population cohort studies? Lessons learned with regards to including and assessing sex, gender and sexual orientation. *Archives of Sexual Behavior* 52, 2163–2172 (Chapter 12)

Ballering, A.V., olde Hartman, T.C., Verheij, R., Rosmalen, J.G.M. (2023) Sex and Gender Differences in Primary Care Help-seeking for Common Somatic Symptoms: a Longitudinal Study. *Scandinavian Journal of Primary Health Care* 41(2): 132-139. (Chapter 7)

Ballering, A.V., Plug, I., Lagro-Janssen, A.L.M., Das, E., Rosmalen, J.G.M. (2022). Building a bridge between multidisciplinary insights and practice: The development of an e-learning for internal residents about sex, gender and persistent somatic symptoms. *Journal of Psychosomatic Research*, 155, 110739 (Chapter 13)

Ballering, A.V., Wardenaar, K.J., olde Hartman, T.C., & Rosmalen, J.G.M. (2022). Female sex and femininity independently associate with common somatic symptom trajectories. *Psychological Medicine*, *52*(11), 2144-2154. (Chapter 5)

Ballering, A.V., van Zon, S.K.R., olde Hartman, T.C., Rosmalen, J.G.M. for the Lifelines Corona Research Initiative. (2022). Persistence of somatic symptoms after COVID-19 in the Netherlands: an observational cohort study. *The Lancet* 400(10350): 452-461. (Chapter 6)

Ballering, A.V., Rosmalen, J.G.M., olde Hartman, T.C. (2022). Differences between women and men are present in the rate of diagnosed diseases after a diagnostic intervention is conducted in primary care. *Journal of the American Board of Family Medicine*. 35(1), 73-84. (Chapter 10)

Ballering, A.V., Oertelt-Prigione, S., olde Hartman, T.C., Rosmalen, J.G.M. (2022) Response to Rossato et al. *Journal of Women's Health* 31(6):896-898 (Chapter 11A)

Ballering, A.V., Ori, A.P., & Rosmalen, J.G.M. (2021). The association of sex, age and FKBP5 genotype with common somatic symptoms: A replication study in the lifelines cohort study. *Journal of psychosomatic research*, *147*, 110510. (Chapter 2)

Ballering, A.V., Oertelt-Prigione, S., Lifelines Corona Research Initiative, olde Hartman, T.C., Rosmalen, J.G.M. (2021). Sex and Gender-Related Differences in COVID-19 Diagnoses and SARS-CoV-2 Testing Practices During the First Wave of the Pandemic: The Dutch Lifelines COVID-19 Cohort Study. *Journal of Women's Health* 30(12) 1686-1692. (Chapter 11)

Ballering, A.V., Muijres, D., Uijen, A.A., Rosmalen, J.G.M., & olde Hartman, T.C. (2021). Sex differences in the trajectories to diagnosis in patients presenting with common somatic symptoms in primary care: an observational cohort study. *Journal of Psychosomatic Research*, *149*, 110589. (Chapter 9)

Groeneveld, J.M., **Ballering, A.V.**, van Boven, K., Akkermans, R.P., olde Hartman, T.C., & Uijen, A.A. (2020). Sex differences in incidence of respiratory symptoms and management by general practitioners. *Family Practice*, *37*(5), 631-636. (Chapter 8)

Ballering, A.V., Bonvanie, I.J., olde Hartman, T.C., Monden, R., & Rosmalen, J.G.M. (2020). Gender and sex independently associate with common somatic symptoms and lifetime prevalence of chronic disease. *Social Science & Medicine*, *253*, 112968. (Chapter 4)

Not included in this thesis

Van Zon, S.K.R**., Ballering, A.V.**, Brouwer, S. Rosmalen, J.G.M. for the Lifelines Corona Research Initiative (2023) Symptom profiles and their risk factors in patients with post COVID-19 condition: A Dutch longitudinal cohort study. *European Journal of Public Health*, ckad152

Ballering, A.V., Kiewiet, S.L. Bakker, B., olde Hartman, T.C., Rosmalen, J.G.M. (2023) Ervaringen van transgender en genderdiverse personen in de huisartsenzorg: een mixed method onderzoek. *Tijdschrift voor Seksuologie* 47(1):14-25

McGhie-Fraser, B., Lucassen, P., **Ballering, A.V.**, Abma, I., Brouwers, E., van Dulmen, S., olde Hartman, T.C. (2023) Persistent somatic symptom related stigmatization by healthcare professionals: A systematic review of questionnaire measurement instruments. *Journal of Psychosomatic Research* 166, 111161

Ballering, A.V., Muijres, D., Uijen, A.A., Rosmalen, J.G.M., olde Hartman, T.C. (2022) Man-vrouwverschillen in de diagnostische trajecten bij veelvoorkomende lichamelijke klachten in de huisartspraktijk. *Huisarts Wet,* 65(5):10-15

Ko, C., Lucassen, P., van der Linden, B., **Ballering A.V.**, olde Hartman, T.C. (2022). Stigma perceived by patients with functional somatic syndromes and its effect on health outcomes – a systematic review. *Journal of psychosomatic research, 154*, 110715

Ballering, A.V., olde Hartman, T.C., & Rosmalen, J.G.M. (2021). Long COVID-19, persistent somatic symptoms and social stigmatisation. *J Epidemiol Community Health*, *75*(6), 603-604.

Slot, L.M., Wormhoudt, T.A.M., Kwakkenbos, M.J., Wagner, K., **Ballering, A.V.**, Jongejan, A., van Kampen, A.C.M., Guikema, J.E.J., Bende, R.J., & van Noesel, C. J. (2019). De novo gene mutations in normal human memory B cells. *Leukemia*, *33*(5), 1219-1230.

Ballering, A.V., Peters, R.M.H., Waltz, M.M., Arif, M.A., Mishra, C.P., Brakel, W., (2019). Community stigma and desired social distance towards people affected by leprosy in Chandauli District, India. *Lepr Rev*, *90*, 418-432.

Heineke, M.H.*, **Ballering, A.V.***, Jamin, A., Mkaddem, S.B., Monteiro, R.C., & Van Egmond, M. (2017). New insights in the pathogenesis of immunoglobulin A vasculitis (Henoch-Schönlein purpura). *Autoimmunity reviews*, *16*(12), 1246-1253. *Shared first authorship

DANKWOORD

Dankwoord

Dit dankwoord is grotendeels op het spoor tussen Alphen a/d Rijn en Groningen geschreven, tijdens zonsopkomst met ABBA op de koptelefoon. Inmiddels kan ik dit traject dromen: de treinreis waardoor ik 'aan de goede kant van de IJssel' terecht kom, met aan jullie de keuze wat de goede kant is. Mijn promotietraject heb ik zeker niet in mijn eentje doorlopen, ik heb veel hulp, steun en interesse van velen gekregen!

Judith, bedankt dat je me altijd een klein beetje of juist heel ver uit m'n comfortzone wist te laten stappen. Hoewel ik het op het moment zelf niet altijd inzag, heeft dit me ongelofelijk veel gebracht. Daarnaast hebben jouw overzicht, rust en bovenal vertrouwen veel stabiliteit in dit promotietraject gebracht. Je zei het vaak genoeg: "Stapje voor stapje, we komen er wel." En je had -uiteraard- gelijk. Door jouw bevlogenheid, kritische houding en betrokkenheid heb ik me geen fijnere én betere promotor kunnen wensen. Bedankt voor alles, ik heb ontzettend veel van je geleerd. **Tim**, wat heb ik een geluk gehad met jou als copromotor. Jouw enthousiasme, vrolijkheid en vertrouwen hebben mijn promotietraject tot een feestje gemaakt. Jouw opbeurende woorden brachten me altijd weer verder. Bovendien zorgde jij ervoor dat mijn stukken ook relevant waren voor de praktijk: "Goh, maar wát heb ik hier morgen tijdens mijn spreekuur aan?" Jouw positiviteit en energie is ongeëvenaard! Bedankt voor alles. En nu we het Noorderlicht hebben gezien, moeten we misschien ook maar op zoek naar het Zuiderlicht?

Hartelijk dank ook aan de leden van de beoordelingscommissie, **Prof. Dr. Schoevers, Prof. Dr. Scherpbier** en **Prof. Dr. den Ruijter** voor de tijd en moeite die jullie hebben genomen om dit proefschrift te beoordelen.

Peter, bedankt voor alle aanmoedigingen, discussies en kritische vragen de afgelopen jaren; ik waardeer het enorm. **Enny**, dank voor jouw scherpe en kritische vragen. En uiteraard voor de gezelligheid. **Wyke, Sandra, Monique,** bedankt voor jullie inzichten! **Sabine**, bedankt voor jouw betrokkenheid, en wat fijn dat ik bij het Nijmeegse gendercolloquium mocht aansluiten! Ik heb er veel over de complexiteit van gender, geslacht en gezondheid geleerd. Met uiteraard ook veel dank aan **Irene**, **Nicole** en **Linda** voor de kritische discussies, maar ook voor het lachen.

Ilona, m'n partner in crime. Hadden we ooit kunnen voorzien dat we in het theater zouden staan met bad-eendjes, glitters en een lichtgevende wc-plopper? Wat ben ik blij dat wij samen ons promotietraject doorliepen en wat heb ik een mooi voorbeeld aan

jou. Bedankt dat je voor me klaarstaat, kritisch op me bent, en de tijd voor me neemt. Ik ben blij dat je naast me staat straks, want jij weet als geen ander mijn zenuwen weg te kletsen. **Tom**, uiteraard ook dank voor alle interesse, het geklets en de lessen over port.

Elske, wat een feest was het de afgelopen jaren. Of we nu Sinterklaasgedichten of papers schreven, het was altijd gezellig. Hoe jij je hand er niet voor omdraait om een evenement uit de grond te stampen, is iets waar ik met veel bewondering naar kijk. Maar ook jouw kritische manier van denken en reflecteren is zeer bewonderenswaardig. Ik waardeer onze koppen koffie samen enorm. Bedankt voor alle steun, het luisterende oor en de goede discussies afgelopen jaren.

Sterre, jouw nuchtere, kritische blik op de wetenschap, en positieve kijk op het leven zijn verfrissend. En je enthousiasme voor het Feel It onderzoek is inspirerend! Je staat altijd voor me klaar en dat waardeer ik ontzettend. Bedankt!

De Feel It! groep met Donald, Sarah en Marijn. **Donald**, hoe jij de rust weet te bewaren, en orde in de chaos van kwalitatieve data weet te scheppen vind ik ongelofelijk. Ik heb heel veel van je geleerd en ook heel hard met je gelachen. Een topcombinatie! **Sarah**, de discussies met jou over gender hebben mijn denken aangescherpt. Ik ben je heel dankbaar dat je altijd even de tijd nam om me aan te horen en met me mee te denken. **Marijn**, jouw achtergrond is heel anders dan die van mij, maar doordat jij me geduldig alles uitlegde heb ik veel van je kunnen leren.

Brodie, wat heb ik genoten van onze discussies en van het werken aan stigma, niet alleen omdat het heel interessant is, maar ook omdat het gewoon heel gezellig is. **Nick**, dank voor alle discussies, of het nu over het wonen in het buitenland, de wetenschap of reizen ging!

Esther en **Margo**, jullie wisten altijd het antwoord op al mijn vragen én jullie vroegen altijd even hoe het met mij ging. Dank voor al jullie geduld en vriendelijke woorden.

Olga, it has been a joy to share an office with you even though I was there infrequently. It was the perfect balance between fun and work (although the scale slightly tips towards fun). **Angélica**, thanks for all the caring words!

Sander, bedankt voor alle aanmoediging, maar ook voor de inhoudelijke discussies. En uiteraard ook dank voor de af en toe broodnodige afleiding in de vorm van koppen koffie, een biertje en bitterballen! Het was fijn om te weten dat jouw deur altijd openstaat. **Inge**, dank voor de nuchtere adviezen, goede raad en het harde lachen.

Daan, **Nord**, en **Els**, wat een goede studenten waren jullie! Het was leuk om met jullie samen te werken, met een aantal gave publicaties als resultaat. **Lucette**, in het eerste jaar van mijn promotietraject ben jij begonnen met een onderzoek in het kader van jouw opleiding. Je hebt m'n promotietraject van begin af aan meegemaakt en was altijd heel geïnteresseerd. Jouw klinische ervaring en kritische houding waren voor mij zeer waardevol!

Het Groene Hart dreamteam, de meesten van jullie ken ik inmiddels ruim 18 jaar... Wat een luxe om te weten dat ik onvoorwaardelijk op jullie terug kan vallen. **Martijn**, wat ben ik blij dat ik jou één van mijn beste vrienden kan noemen! Bedankt voor alles wat je voor me hebt betekent de afgelopen jaren. De vele potten thee worden ook zeer gewaardeerd! **Esther**, bedankt voor alle interesse en aanmoediging de afgelopen tijd. Ik ben onder de indruk van hoe jij in het leven staat. Je bent een topper, zet 'm op de aankomende jaren! **Miles** en **Emma**, bedankt voor alle interesse, kritische vragen en lieve woorden. **Robert** en **Pascalle**, jullie zijn de nuchterheid zelve en altijd bereid m'n Groningse verhalen aan te horen en vervolgens met goede raad te komen. Dank je! **Simon** en **Marie**, dank voor het alle interesse! De kritische (soms filosofische) vragen die me gesteld werden hebben mijn denken aangescherpt.

Esther, **Mirjam**, **Dewi** en **Irene**: áltijd geïnteresseerd, áltijd in voor een pittige discussie, maar bovenal áltijd een luisterend oor gevolgd door waardevolle adviezen (of soms een arm om m'n schouder). Bedankt! **Wendy, Menno, Sharon** en **Marc**, het is een luxe te weten dat jullie altijd voor me klaarstaan, met een kop thee, lekker eten, een goede plaat en Jip als knuffelhond.

Alle heren van de **BDCG** en de **aanhanger**-dames: bedankt voor alle interesse de afgelopen jaren. In het bijzonder **Niels, Nicole, Maikel** en **Mirjam**, bedankt voor alle aanmoediging, kritische vragen en leuke discussies de afgelopen jaren.

Oma en **Tante Co** het is zover! Ik ben nu écht bijna klaar. Bedankt voor alle geborgenheid, warme woorden en steun.

Pap, bedankt voor alle onvoorwaardelijke steun, maar ook afleiding door de jaren heen. Want we hebben wel gelachen tijdens ons wekelijks rondje hardlopen! Heel snel zijn we niet, maar ja, wat wil je met al die koffie tussendoor? **Désirée**, bedankt dat ik altijd mag binnenvallen voor een kop thee (met minstens 2 koekjes erbij; immers, op één been kun je niet lopen) en dat je het zag als het even niet lekker ging. Ik ben blij dat jullie er altijd voor me zijn.

Kirsten en **Kevin**, bedankt voor alle nuchtere adviezen en hulp de afgelopen jaren. Jullie geruststellende woorden en dikke knuffels waren soms het enige wat ik even nodig had. Ik kan me geen fijnere zus en zwager voorstellen!

Mark, zonder jou was de afgelopen tijd lang niet zo leuk geweest. Niet alleen ben ik je dankbaar voor alle ritjes naar het station (of *all the way* naar Groningen), de ontelbare kopjes koffie en emmers thee die je me in de loop der jaren hebt gebracht (mét koekjes uiteraard). Ik ben je ook dankbaar voor al je geduld, liefde en steun. En zelfs voor alle keren dat je mij zonder enige moeite op m'n plek zette als het nodig was (lees: als jij het nodig vond). Het is ongelofelijk hoe jij de rust wist te bewaren als ik die kwijt was. Ik heb er heel veel zin in om de aankomende tijd met jou te verdwalen de toekomst in. Kom op, ik weet de weg!

En dan tot slot, **Evelien** en **Mariska**, bedankt voor het meedenken in het vormgeven van dit proefschrift. Jullie creativiteit gaat mij de pet te boven!

RESEARCH INSTITUTE SHARE

Research Institute SHARE

This thesis is published within the **Research Institute SHARE** (Science in Healthy Ageing and healthcaRE) of the University Medical Center Groningen / University of Groningen.

Further information regarding the institute and its research can be obtained from our internet site: https://umcgresearch.org/w/share

More recent theses can be found in the list below (supervisors are between brackets).

2023

Boersema HJM
The concept of 'Inability to Work Fulltime' in work disability benefit assessment
(Prof S Brouwer, Dr FI Abma, Dr T Hoekstra)

Ots P
The role of individual and contextual factors in paid employment of workers with a chronic disease
(Prof S Brouwer, Dr SKR van Zon)

Kool E
Untangling the elements of midwives' occupational wellbeing: A study among newly qualified and experienced midwives
(Prof ADC Jaarsma, Prof FG Schellevis, Dr EI Feijen-de Jong)

Jansma FFI
Self-management in rehabilitation practice: On the design and implementation of a serious theory-based analogue problem-solving game called 'Think Along?'
(Prof R Sanderman, Dr I Wenzler)

Erpecum CPL van
The role of fast-food outlet exposure in Body Mass Index
(Dr N Smidt, Prof U Bültmann, Dr SKR van Zon)

Kerver N
The effectiveness and cost-effectiveness of upper limb prostheses
(Prof CK van der Sluis, Dr RM Bongers, Dr S van Twillert)

Deviandri R
Management of anterior cruciate ligament injury in lower-middle income countries: Focus on outcomes and health economics in Indonesia
(Dr I van den Akker-Scheek, Prof MJ Postma, Dr HC van der Veen, Dr Andri MT Lubis)

Mangot Sala L
Disruptive Life Events and Health: Longitudinal evidence from a large cohort in the Netherlands
(Prof AC Liefbroer, Dr N Smidt)

Wijk DC
From prosperity to parenthood: How employment, income, and perceived economic uncertainty influence family formation
(Prof AC Liefbroer, Prof HAG de Valk)

For earlier theses visit the website: https://research.rug.nl/en/publications

Research Institute
SHARE