Doctoral theses at NTNU, 2022:157

Ellen Rabben Svedahl

Health services under pressure:

General practitioners' workload, gatekeeping, and patient safety

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Medicine and Health Sciences
Department of Public Health and Nursing

NTNU
Norwegian University of
Science and Technology

Ellen Rabben Svedahl

Health services under pressure:

General practitioners' workload, gatekeeping, and patient safety

Thesis for the Degree of Philosophiae Doctor

Trondheim, May 2022

Norwegian University of Science and Technology
Faculty of Medicine and Health Sciences
Department of Public Health and Nursing

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Medicine and Health Sciences
Department of Public Health and Nursing

ISBN 978-82-326-5430-7 (printed ver.)
ISBN 978-82-326-5841-1 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2022:157

Printed by NTNU Grafisk senter

Sammendrag

Bakgrunn

Med en økende og aldrende befolkning verden over øker presset på helsevesenet og kapasiteten utfordres. Helsevesenet i Norge er rangert som et av verdens beste, og har en universell utforming som skal sikre at alle som trenger helsehjelp får det de har behov for. Ideelt sett skal pasienter behandles på laveste effektive omsorgsnivå, og ulike tiltak settes stadig inn for å begrense kostnader og ressursbruk. Tidligere forskning har imidlertid vist betydelige variasjoner i helsetjenestebruk, som kan tolkes som uttrykk for både skjevfordeling og mulig overforbruk. Slik variasjon, blant annet i henvisningspraksis, er vist å ha sammenheng med ulike egenskaper ved leger og ved helsetjenesten. Dette er forhold som er vanskelig å undersøke på en god måte; randomiserte kontrollerte studier er etisk og praktisk utfordrende, mens observasjonelle data er sårbare for skjevheter, ved at tjenestene og pasientgruppene ikke er sammenliknbare. Dette preger den eksisterende litteraturen. Konsekvensene av slik variasjon er også sparsomt beskrevet tidligere. Med stadig økende behov for helsetjenester, trengs det videre målrettet forskning på dette viktige området for å sikre bærekraftige helse- og velferdstjenester av høy kvalitet.

Formål

Formålet med forskningen som presenteres i denne avhandlingen var å undersøke hvordan ulike faktorer påvirker dynamikken og særlig pasientflyten mellom primær- og spesialisthelsetjenesten, og hvordan dette igjen påvirker videre helsetjenestebruk og pasientsikkerhet. Vi ønsket å undersøke hvordan arbeidspress påvirket måten legene arbeider på, hvordan ulike legeegenskaper påvirker pasienters sjanse for å bli lagt inn på sykehus og deres påfølgende helsetjenestebruk. Videre ønsket vi å undersøke hvordan eldre pasienter på legevakt blir påvirket av legens beslutning om henvisning til sykehus, særlig i de tilfellene hvor det kan være tvil om henvisningen. I tillegg til å se på pasientsikkerhet, ønsket vi også å belyse hvordan beslutningen om henvisning kan tenkes å belaste eller avlaste helsetjenesten gjennom endret bruk av fastlege eller spesialisthelsetjenesten for disse pasientene i tiden etter legevaktkontakten.

Metode og materiale

I denne avhandlingen presenteres bruk av flere metoder i tilnærmingen til forskningsspørsmålene. Med kvalitative intervju har vi undersøkt hvordan fastleger og deres medarbeidere opplevde arbeidsbelastning i hverdagen sin. Vi intervjuet 23 norske fastleger og 10 av deres medarbeidere, noen i individuelle intervju, noen i fokusgrupper. Intervjuene ble tatt opp, transkribert og analysert ved hjelp av systematisk tekstkondensering (artikkel 1). I de kvantitative analysene av store norske registerdata har vi forsøkt å ta hensyn til utfordringer som ofte følger analyser av observasjonelle data. Fortolkningen av kvantitative analyser av registerdata må ta høyde for betydelige utfordringer sammenlignet med intervensjonsstudier. Vi har forsøkt å velge analytiske forskningsdesign som adresserer flere av utfordringene som analyser av slike registerdata har. Selv om målet er å komme nærmere begrunnede årsakssammenhenger, må vi huske på at fortolkningen av resultatene fra analysene hviler på noen avvgjørende forutsetninger om sammenliknbarhet. Ved å velge ut og sammenlikne grupper av pasienter som i teorien skulle ha like forutsetninger for innleggelse, forsøkte vi å anslå effekten av ulike legeegenskaper på deres sjanse for innleggelse (artikkel 2). Videre brukte vi kunnskapen om legevaktslegenes varierende terskel for å henvise pasienter til innleggelse, til å modellere en situasjon hvor denne egenskapen, som vi kalte legens *henvisningspreferanse*, tilfeldig påvirket ellers like pasienters sjanse til innleggelse. Her valgte vi ut pasienter som vi kunne gå ut fra at var ukjente for legen og legevakten, ved at de ikke hadde vært vurdert av samme lege på legevakt i perioden, og ikke var såkalte storforbrukere av legevakt. Basert på dette kunne vi anslå effekten en slik innleggelse har for pasientens videre helsetjenestebruk og risiko for å dø (artikkel 3).

Resultater

I vår kvalitative studie (artikkel 1) fant vi at fastlegene og medarbeiderne vi intervjuet opplevde et stadig økende arbeidspress i hverdagen sin, i tillegg til naturlige svingninger gjennom uken og året. De pekte på et vidt spenn av årsaker til arbeidspress, og det var stor variasjon både blant legene på hvert kontor og mellom kontorene på hva de anså som de viktigste årsakene. Videre var de oppmerksomme på at høyt arbeidspress hadde

hovedsakelig ugunstig påvirkning på måten de arbeidet på, og de uttrykte bekymring for at dette ikke var bærekraftig. De var bevisste på sin portvokterfunksjon, og beskrev at de bevisst forsøkte å unngå økning i henvisninger, da dette kunne føre til merarbeid både for dem selv og helsetjenesten. I den første kvantitative studien (artikkel 2) fant vi at ulike egenskaper hos fastleger som jobber legevakt, har betydelig påvirkning på pasientenes sjanse for å bli akuttinnlagt. Mannlige leger, og eldre leger, la inn en lavere andel av pasientene enn sine kvinnelige og yngre kolleger. Det var liten forskjell på om legen var spesialist i allmennmedisin eller ikke, mens leger med lav erfaring la inn en litt høyere andel pasienter. I møte med like pasienter var det som så ut til best å kunne forklare legens beslutning om innleggelse, denne legens *innleggelsespraksis* de siste fire måneder, målt som andel av denne legens tidligere pasienter som ble akuttinnlagt fra legevakt. I denne studien fant vi små forskjeller i dødelighet de neste 30 dager etter en legevaktkontakt, knyttet til egenskaper ved legevaktslegen. I den neste kvantitative studien (artikkel 3) fant vi at pasienter over 64 år som vi antok var ukjente for legen, og som fikk henvisning til sykehus *basert på legens henvisningspreferanse*, hadde betydelig mer sykehusbruk de første ti dagene, sammenliknet med å ikke henvises. Videre fant vi at disse pasientene også fikk høyere bruk av spesialisthelsetjenester og også etter hvert primærlegetjenester i løpet av det påfølgende halve året. Et viktig funn var at pasientene som ble anslått å ha blitt henvist på bakgrunn av legens henvisningspreferanse også hadde omtrent halvert risiko for å dø de første 10 dagene etter legevaktkontakten.

Konklusjon

Funnene som presenteres i de tre artiklene er tilskudd til kunnskapsgrunnlaget som trengs for å kunne opprettholde og bedre kvaliteten i helsetjenesten. Fastlegene opplever at arbeidsbelastningen påvirker deres måte å jobbe på, og de er bekymret for økt gjennomtrekk og vansker med rekruttering. Dette er i samsvar med de endringene vi også ellers ser i fastlegetjenesten, og vil kunne påvirke sammensetningen av legene som jobber i legevaktstjenesten. Dette vil igjen, gjennom blant annet effekter på portvokterfunksjon, kunne påvirke akuttinnleggelser, videre helsetjenestebruk og sikkerhet for pasientene, som i sin tur fører til økt belastning på helsetjenesten. Basert på våre funn, ser det ut til å kunne være lønnsomt for både primær- og

spesialisthelsetjenesten å gjøre tiltak for å bevare og styrke primærhelsetjenesten og portvokterfunksjonen. Større tiltak bør imidlertid evalueres gjennom forsøk før eller samtidig med at de settes i verk.

Ellen Rabben Svedahl
Institutt for samfunnsmedisin og sykepleie, NTNU

Hovedveileder: Johan Håkon Bjørngaard
Biveiledere: Kristine Pape og Bjarne Austad
Finansieringskilde: Norges forskningsråd – HELSEVEL.

In loving memory of Ingrid Johansen Aune,
who always believed in working to make the world a better place for all.

Acknowledgements

This PhD project is part of the research project "Pressure in the health care systems, consequences for efficiency, patient flow and patient safety" funded by the Norwegian Research Council (grant number 256579). It was also supported by the Department of Public Health and Nursing at the Faculty of Medicine and Health Sciences, NTNU. I highly appreciate the opportunity and funding allowing me to do this work.

Through working with understanding causal relationships, I have gained an even greater perception of the influence of chance. Being born in Norway, given the opportunity to study and practice medicine in what may be the best health services in the world is unfortunately a privilege for the few. Although I had always wanted to get into research, it was by chance that I came across this PhD position and could suddenly learn and perform research on this theme that had been on my mind for a long time. I feel so fortunate to be included in the inspiring research environment at ISM, getting to know new colleagues and gaining friends for life. I am ever grateful for the support from all who have contributed to my work with this thesis.

There are some people to whom I would like to express my special appreciation. First and foremost, I would like to extend my deepest gratitude to my main supervisor Johan Håkon Bjørngaard, for accepting me for the PhD position and supporting me through this period. Thank you for sharing your immense knowledge, numerous ideas, critical advice, and for your seemingly endless patience with me (even after sleepless nights where I barely remembered which day it was). Your enthusiasm, comprehensive methodological insights, creativity, and literature tips have been decisive for the work with this thesis, my understanding of epidemiological research and for developing my views and perspectives on medicine and health services research.

I also give my warmest thanks to my co-supervisor Kristine Pape, for including me from the very first day, always being kind, caring and supportive, and for valuable critical advice making me think through the meaning of what I have written. Your Stata skills and register data experience have contributed substantially to my project. Together with your talented daughter Juni Pape Skjærvold, you have also made unique and illustrative figures that have been very useful for disseminating our research. Thanks to both of you. Further, I would like to thank my co-supervisor Bjarne Austad, for your support, critical input, and for helping me put everything in a relevant perspective. Your broad experience from general practice and research have been highly valuable for my project.

I highly appreciate being included in the project group and participating in the stimulating research environment, benefiting from all the different perspectives, and

getting to know all the inspiring people in the group. Kjartan, thanks for giving fast and invaluable support on our data, Stata and other computer endeavours. You have saved my day so many times! Further, my co-authors Marlen, Lena and Silje, you have kindly shared your knowledge and experience, tips, tricks and life-hacks, included me in memorable field trips, and generally made research a whole lot of fun. My deepest thanks to my co-authors in the quantitative papers, Gunnhild, Fredrik, and Neil, for valuable critical revisions of ideas and manuscripts. And thanks to the whole Regforsk-group, which I am very proud to be part of, especially to Sara, Andreas, Gudrun, Christina and Kristine V, for your inspiring support and collaboration.

Further, I would like to thank Epinor and my fellow students for inspiring meetings and challenging discussions, developing our understanding of epidemiology, and working with the 24/7 project, which provided valuable research experience. Special thanks to Katja; you have been an invaluable support to me personally and professionally.
Little did I know that accepting the PhD position would also include me in a unique office team – Team Corner Office. Thank you, Vegard and Maria, for being the special persons you are; I am deeply grateful for your support with research - and everything else in life. And for all the fine chocolate. I was last in and last out, but it sure has been a pleasure making this journey with you.

I would also like to thank all my good colleagues on the 4th Floor, including Sindre P, Unni, Christine, Melanie, Marion, Pål, Bjørg, Signe, Kirsti, Kjersti, Nils Arne and Torunn, for always making the office an enjoyable place to be, for all your support, cheering and exciting coffee and lunch discussions. Also, thanks to fellow researchers and staff on other floors at the ISM for valuable feedback, support, and help. It has been a pleasure working with all of you, and I look forward to the continuation.

Further, I would like to express my gratitude to the Department of Dermatology at St. Olavs hospital for providing a stimulating work environment, allowing me to do research, and to all my good colleagues there for supporting me through this process. Working with you has given me valuable experience and perspectives, and I look forward to returning to the clinic.

I am also immensely grateful for the backing and encouragement from my highly valued friends, and for having such a wonderful and supportive family. My sister, Elisabeth, thank you for your inspiration and encouragement and always caring for our family and me. The warmest thanks also to my parents, Else and Ulrik. Through your loving care, you have raised us to believe that we can be what we want to be and do what we want to do, while constantly remining us that we are always good enough just the way we are. I hope I can pass this on to our children. I have also been so lucky to have Anne Marit and Stein, my parents-in-law, in my life. You kindly included me in

your family when I was 16 and have been stuck with me since. Thank you for your endless support, love, advice, and inspiration. Thanks to all our children's grandparents for your understanding and practical help with running and managing the family life. This PhD would never have been possible without you.

So, to our wonderful girls, Stine Elisabeth, Ane Linnea, Ingrid Elea and Nora Leannah. You are my greatest inspiration for working to make the world a better place and believing in the future. Thank you for always being there, sharing your curiosity, questions, stories, unconditional love and support. Playing, singing, dancing, laughing, talking and just being with you makes everything else bearable. And last, and most importantly, my husband, Sindre. You are the wisest person I know and my inspiration in life. You can always make me laugh. Thank you for always loving me, believing in me, supporting me, and making me want to be the best version of myself. And for being the best dad for our girls. With you, I feel complete.

Malvik, April 2022
Ellen Rabben Svedahl

List of papers

Three papers are included in the thesis and will be referred to as Paper I, II and III.

Paper I:

Workload in general practice – a qualitative study from Norway

Ellen Rabben Svedahl, Kristine Pape, Marlen Toch-Marquardt, Lena Janita Skarshaug, Silje-Lill Kaspersen, Johan Håkon Bjørngaard, Bjarne Austad

Published in BMC Family Practice: 21 May 2019
Doi: https://doi.org/10.1186/s12875-019-0952-5

Paper II:

Effects of GP characteristics on unplanned hospital admissions and patient safety. A nine-year follow up of all Norwegian out-of-hours contacts.

Ellen Rabben Svedahl, Kristine Pape, Bjarne Austad, Gunnhild Aaberge Vie, Kjartan Sarheim Anthun, Fredrik Carlsen, Johan Håkon Bjørngaard

Published in Family Practice: 2021
Doi: https://doi.org/10.1093/fampra/cmab120

Paper III:

Out-of-hours referral to hospital for admission – The impact of altering the referral threshold on patient safety and further health service use

Ellen Rabben Svedahl, Kristine Pape, Bjarne Austad, Gunnhild Aaberge Vie, Kjartan Sarheim Anthun, Fredrik Carlsen, Neil Martin Davies, Johan Håkon Bjørngaard

Ready for submission

Abbreviations and acronyms

GP	General Practitioner
RGP	Regular General Practitioner
OOH	Out-of-hours
RCT	Randomized Controlled Trial
DAG	Directed Acyclic Graph
OR	Odds Ratio
CI	Confidence Interval
ATE	Average Treatment Effect
LATE	Local Average Treatment Effect
IV	Instrumental Variable
DRG	Diagnose Related Group
PCP	Primary care physician (including both physicians working as GPs during daytime and out-of-hours).
ICD-10	International Statistical Classification of Diseases and Related Health Problems
ICPC-2	The International Classification of Primary Care

Table of contents

1 Preface

On my adventure through clinical practice in medical school, as an intern in a smaller hospital, GP locum in primary care and out-of-hours service, and then specialising physician in a large university hospital, I have observed and reflected over both patients´ flow through the health care system, and health care staff making different decisions on behalf of the patients. I have also observed, and later also experienced, that these variations in decisions may be affected by many different factors; personality, education, personal and professional experience, workplace culture, recent events at work, at home or in the media, how the day has been so far, and how the work schedule is planned, to mention some. And importantly, also the perception of the patients´ expectations. I guess most patients have high expectations of fair treatment, and in Norway, our health system is based on universal access, to secure that all patients get the care they need. This should be a fact regardless of the physician they meet, at least within an acceptable level of variation.

But what happens when the framework for performing care is compromised? When there are too many decisions to make, in too short time? When the capacity limits are reached, the stress level is too high, there is too little competence or experience, or limited opportunities to seek help from more experienced or colleagues? I have also observed these factors to influence decision making. It might feel easier to escalate the patients to a higher level of care, just to be on the safe side. And on an individual level, this might seem like or even be the right thing to do. But how will these decisions affect the patient in the long term? And how will they affect the health services, and other patients on a system level? I imagine that the consequences of all these small decisions in sum are much larger than one might think. Knowledge about how the size of variation, factors affecting it, and not least, the consequences of this variation, is essential for assessing the need for action, and also to understand which actions will have the desired effects.

I´ve pondered these questions quite a lot. However, I realised that my number of observations was small, and likely biased by my own preconceptions. When I came

across the project: “Health care services under pressure, - Consequences for patient flows, efficiency and patient safety”, I immediately saw the potential to both obtain competence as a researcher in an important field, and to potentially contribute with valuable knowledge, for improving the framework for the health care providers. When accepting the position as PhD candidate on the project, I prepared to go into heavy register-based research, assuming my days would consist av handling large numbers and statistical computer programs, buried in statistical books. However, my days have comprised interesting reasoning and discussion about the actual field knowledge needed before even looking at the numbers. I felt my previous education and experience was useful for contributing to planning good research, and that my input was valued. Fortunately, I must say in hindsight, we had to wait for the data to arrive. During this time, I got involved in a qualitative study, which became part of my project. The aim was to generate good research questions for the quantitative analyses (because both my supervisors, Johan Håkon and Kristine, acknowledged the benefits of knowing the field thoroughly before asking research questions). I was tempted by the idea of also getting experience with qualitative methods during my PhD period, and I am very happy about accepting the offer to write a qualitative paper first. This also led to adding another co-supervisor, Bjarne, who has contributed substantially to my understanding of qualitative research. This sub-project provided valuable insight into the field of the quantitative analyses. Also worth mentioning, I achieved a more nuanced perspective of time, and we had our fourth daughter.

And about trust.

In health care as well as research, trust is a key element. For relying on research results, it is important to have trust in both the data you put into the analyses, the methods used to analyse the data, and to have trust in the researcher and his or her intentions for doing the research. I hope my background as a physician, in both primary and secondary care, but also as a patient and next of kin (though with n=a minimum), will contribute positively to my work, however I also acknowledge my preconceptions from these experiences. Hopefully, my robust research environment, with highly experienced and qualified supervisors and colleagues has helped me in handling these challenges. Our work is based on comprehensive high-quality data from the Norwegian health care

system. Further, the combination of both a qualitative approach to the research field, and carefully planned and adjusted quantitative methods, may enhance the trust in the use of observational data for providing useful and reliable knowledge in health services research. Trust is also essential in the intersection between patients and physicians, both to optimize the perception and the treatment, but especially for limiting the use of health services in cases where more is not better. (1)

2 Background

The current pandemic has imposed an extreme pressure on the health services across the world, and capacity limits are reached or even exceeded. Even before the pandemic hit in 2020, the increasing demands from an ageing population were evident. In Norway, as in many other developed countries, the population is both increasing and ageing, more people live with chronic illness, and treatment opportunities and medical technologies are rapidly expanding. At the same time, the workforce is shrinking. Nevertheless, public expectations persist or even increase, both for the quality of health services and for health in general.

The Norwegian health care system is repeatedly ranked as one of the best globally (2) (3), with high scores on accessibility and patient satisfaction and high survival rates for specific diagnoses. Concurrently, our health care system is one of the most expensive in Europe. In 2017 health care expenditure accounted for 10.4% of Gross Domestic Product, placing Norway fifth highest in the WHO European Region. (4)

With the aim of universal access to high-quality care, preserving and improving the Norwegian health care system has been high on the political agenda. In the latest decades, policymakers have become increasingly aware of the challenges from rising demands. This thesis is based on a project established in 2017 exploring pressure on capacity limits in the health care system and the consequences for dynamics, patient flow, and patient safety. Even with increasing emphasis on preventing diseases, the demands for health services are steadily rising, reflected in unsustainable increases in waiting lists and health services expenditure, especially from specialised health services. (5)

In March 2020, the health care sector became the centre of attention, as the pandemic accentuated the visibility of the current capacity challenges in the health services. Further, it underscored the need for universal access to high-quality health services, both to prevent severe illness and limit virus spread. Lessons learned after almost two years on alert are that the spare capacity was relatively scarce even before the pandemic, and that increased pressures in some parts of the services inevitably affect the whole

system. Consequently, there is a need for profound organisational changes to avoid a breakdown, which is already underway in several European countries.

Reducing pressure and improving the cost-benefit ratio from health services expenditure can be achieved in several ways. In addition to temporarily increasing efficiency within each part of the service by running faster and working smarter, there is a need for more long-term interventions. Moving specific tasks and responsibilities to less costly and resource-demanding services is one possible approach such as transferring rehabilitation and follow-ups for chronic disease from specialised services to primary care. Also, shifts from inpatient to outpatient treatment and surgery and discharging patients earlier will potentially relieve hospital costs. These measures, among others, were formalized in Norway through the implementation of the Coordination reform in 2012. (6) The reform also intended to strengthen primary care by increasing the number of General Practitioners (GPs) accordingly. However, the modest increase in the number of GPs is arguably not in line with the population demand, imposing an added workload in primary care. (7)

Another potential measure that has been debated is reducing the inflow of patients to the specialised services. In Norway, as in many European countries, access to secondary care is regulated through a gatekeeper function, requiring assessment and referral by a physician, unless there is an emergency condition. The indications for referral are currently regulated through the Norwegian Specialist Health Service Act (8) and aided by the National guidelines for referrals to the specialist health service (9). However, there is room for individual interpretation/judgment, and substantial variations in both elective and acute referral rates suggest that there is potential for optimisation of referral practice. (10-12)

To be prepared for what is to come, and adjust accordingly, knowledge about the current system's dynamics and complexity is essential. A starting point is to evaluate the current situation and identify weak points and potential targets for improvement. Further, there is a need to evaluate the effect of changes and interventions. These evaluations require accurate descriptions of the system and reliable methods to assess

the causal relationships between modifiable factors and various outcomes. However, defining adequate measures for modifiable factors and outcomes in health services is difficult. Comparisons are often used to determine superiority; however, causal inference from observed associations is difficult, since subjects of interest are seldom comparable. Finding good ways of comparing patient outcomes from the different organisational models between different countries, regions or even different parts of the health care sector imposes a great challenge because of potentially confounding factors, such as systematically differing populations. Combining resources from different research fields, traditions and methods can enlighten different aspects of the problem. There is also increasing interest in the use of register data for this purpose, and there are plans for providing such data for more extensive parts of the health service, also in the other Nordic countries. (13)

Nevertheless, by puzzling small pieces of valid evidence together, the sum can become a valuable contribution to understanding the system, and for evaluating the effect of potential or implemented measures.

In this thesis I will address some essential aspects of how pressure on the health services might affect the dynamics, patient flow and patient safety in the Norwegian health care system. All three papers are situated at the intersection between primary care and specialised services, and the focus in this thesis will therefore be on these two parts of the system and the dynamics between them. Since the health services are in constant transition, I have supplemented my background section with contemporary literature. I will start by providing an overview of the Norwegian health care system partly in relation to in an international context. Further, I will describe more closely the key aspects of primary care and specialised services in Norway. In the description of the general practitioner service, I will enlighten aspects from both the regular general practitioner service (in normal working hours) and the out-of-hours service and explain the gatekeeper role. Further I will address variations in health care use, and different views on this topic, including the use of variation in patient outcomes for quality assessment in health services, and the challenges related to overuse of health services without benefits for the patients or the society. I will also give a short overview of

workload in general practice and the out-of-hours services and how increased pressure might affect the dynamics of the health care system.

2.1 The Norwegian Health Care system

In 2021, Norway´s population will comprise almost 5.5 million inhabitants. The Norwegian Health Care system is inspired by the English National Health Service (NHS). Like the other Nordic health care systems, it is mainly publicly funded and founded on the principles of universal and equitable access, regardless of geography, socioeconomic status or ethnicity, and aims to offer comprehensive and high-quality health services. (14) In 2021, there are 5.0 physicians and 17.9 nurses per 1000 inhabitants, which is among the highest in Europe. (OECD 2021)

There are two main operational levels; primary care and secondary care.
In primary care, accessibility and proximity to the patients are prioritised, and the services are intended to be comprehensive in offering support in all stages of life. Primary care is financed and organised by the municipalities and comprises GP services, out-of-hours services, rehabilitation, physiotherapy, psychology, mid-wife services, nursing stations and care homes. Secondary care offers specialised health services for investigations, treatments and follow-ups of patients with acute, severe and chronic illness and health conditions, where the needs exceed what can be covered by primary care. (8) It includes (but is not limited to) hospitals and specialist outpatient clinics, laboratory and medical imaging services, psychiatric services, rehabilitation services, drug addiction institutions and some prehospital services.

Secondary care is organised in two administrative levels, the national and regional levels, both funded by the state. The regional level is divided into four Regional Health Authorities, responsible for the 20 hospital trusts. (15) Notably, this implies that primary and secondary care are financed over different budgets. Patients in primary care are paid for by the municipalities, while the state pays for patients in secondary care. Hence, there is no financial incentive to prevent referring patients to specialised health

care, other than the daily fines/co-payment stated to primary care for patients ready to be discharged from the hospital. (16)

Based on the aim of equitable access, there are limited out-of-pocket payments for health services, with an annual cap ceiling of approximately 275 $ per patient (2021) for GP services, specialist outpatient clinic visits, as well as for selected prescription drugs and medical equipment, psychologist and physiotherapist services, and patient travel expenses. (17) Adult dental care and care homes are not included in these services and are still mainly paid for by the users. (15) For children under the age of 16, there are no co-payments for any of these services.

Access to secondary care is regulated through the gatekeeper function, and hospital admissions and ambulance services are free of charge for all patients.

The Nordic countries are all based on the same welfare model and have many similarities. They are organised in primary and secondary care and have high shares of public funding. Although they aim for universal access to care and are generally ranked high in ratings of health care systems, there are some notable differences. In Norway and Denmark, the GPs are primarily self-employed and act as gatekeepers for most specialities. An exemption is that the Danish public can contact ophthalmologists and ear-nose-throat specialists directly. A comparative study from 2017 showed that daytime General Practice in all the Nordic countries were generally highly equipped and provided a wide range of medical procedures. (18)

As Denmark and Norway may have the most similar systems, it is worth noting that the geography of the countries is different, thus requiring somewhat different approaches to secure sufficient provision of services. The English health care system is also quite similar, with the National Health Services securing universal health coverage for all residents. Services are mostly free at the point of use, GPs also engage in gatekeeping, and the hospitals are mostly public. However, the demography and geographical setting differ even more than from the Nordic countries. (19)

2.2 General Practice

Primary health care services, including General Practice, are often regarded as the cornerstone for the health care system in Norway and many other countries. (20) Primary care consists of several actors; however, I will mainly emphasise the General Practitioner service and especially the out-of-hours services for this thesis.

2.2.1 The regular GP service

Based on a need to better coordinate services and facilitate continuity of care, the Regular GP service was established in 2001, after successful pilot projects in selected municipalities from 1995. (21) From 2001 all inhabitants in Norway were invited to choose a regular GP. The vast majority of the population accepted this offer, and in 2021 about 98% are affiliated with a regular GPs list. (22) The regular GP service has been shown to be one of the most popular public services in Norway, with high patient satisfaction over the years. (23)

In 2019, the total number of consultations with regular GPs in Norway was approximately 14,8 million, hence a mean number of 2,7 consultations per person. This number has steadily increased since 2015. (24) The share of the people using the regular GP services is higher among women, and also increases with increasing age.

A specialisation in general practice (in the literature also referred to as Family Medicine) was established in 1985. From 2017, taking part in systematic training to obtain a specialisation in general practice became mandatory for all regular GPs, (unless already qualified). (25) This requires supervision both individually and in groups, educational courses, training in other practice fields, such as hospital wards, participation in the out-of-hours services, and at least two years working in general practice.

'The Regulations relating to a Municipal Regular GP scheme' (26) states that the regular GPs are obliged to be available for contacts from their patients during office hours to handle acute and non-acute medical needs, and offer home-visit where this is indicated. It also states the GPs obligation to perform municipal general practitioner

tasks (up to 7,5 hours per week), and to participate in the out-of-hours service provided by the municipalities.

Nine out of 10 GPs work in groups, and most of them are self-employed, with responsibility for managing a GPs office. Payments are a combination of capitation fee (per patient on their list) and consultation fee. In 2018 there were approximately 4,750 GPs in Norway, which increased to 4,930 in 2020 (27). In 2020, 46 per cent of the GPs were female, and 63 per cent were specialists in general practice. The mean number of patients registered pr GP peaked at 1,200 in 2005 and has slowly decreased to 1,068 in 2021 (998 for female GPs, 1,127 for male). (22)

As many of the physicians who joined the regular GP scheme in the early 2000, are still serving their list population, the mean age of regular GPs is relatively high. (22) Thus recruitment of new GPs will probably lead to a lower mean age, and lower mean experience than what is currently the situation. Further, a higher share of female medical students the later years, will most likely also contribute to a higher share of female GPs. (28) This expected change in composition of GPs may influence their practice both as regular GPs and through their participation in the out-of-hours services.

The coordination reform

In the foreword of the whitepaper from 2009 describing the Coordination reform, the current Minister of Health Bjarne Håkon Hanssen, described his vision for the reform, referring to this phrase: "*..the patient is contemporarily discharged to the specialised health services..",*emphasising that the base of the patient care should be in primary care, and not in specialised care, and implying that this is not usually the case. (6)

The Coordination reform was implemented in 2012 to improve the coordination and cooperation between primary and secondary health care. (6) To new acts described the new structure; The Norwegian Public Health Act (29) and The Act of Municipal Health and Care Services. (30) Primary care was given the responsibility for more tasks to relieve the specialised health services from increasing pressure, to prevent fragmenting

of services, strengthen organised preventive health care, and to be able to adapt to the expected demographical change.

Primary care was strengthened through earmarked grants to stimulate municipal investments in Municipal Acute Wards as an alternative to costly hospital admissions. However, there were also two additional financial incentives for limiting the municipalities use of specialised health care, namely and municipal payments for patients ready for discharge and co-financing of hospital treatments for resident patients. The latter was dropped after two years, as it showed no substantial effects on hospital admissions. (31)

The reform underscored the importance of the GP role as a medical coordinator for the patient. However, it also emphasized the GP´s function as a gatekeeper for specialised health services. It described a need for strengthening the GP´s assessment competence, to reduce unnecessary referrals by the following statement: "*When GPs are to contribute to fewer - ie the right patients - being referred to hospital, the GP's competence must also be strengthened."* To obtain this, both the assessment competence and the clinical competence for treatment needed reinforcement. The gatekeeper function is described in more detail in section 2.1.5.

Because of the expected rise in workload per patient, due to both the reform, and the ageing and increasing population, there was an intention of downsizing the patient lists accordingly by recruiting more physicians into General Practice. In the almost ten years since the reform implementation, the number of GPs has increased, but so has the population. Thus, as the regular GP practices has increased from 4,279 to 4,930, the average number of patients per GP list has decreased from 1,164 in 2012 to 1068 in 2020. Nevertheless, GPs report increasing workload, by both consultation rates (27) and reported working hours. (32)

Despite the vision of strengthened primary care and increasing evidence suggesting substantial advantages from continuity of care in the GP service, many GPs, policymakers, and patients currently express substantial concerns for a fragmentation of

the regular GP scheme. (7, 33-35) The share of patients affiliated with lists without a regular GP is now 3.5 % (representing 2 % of the patients) compared to only 1.6% of the lists in 2012. Almost 22 % of the patient lists were registered as served by a locum in the past year. (22)

Diversity in General Practice

There are some key features in how GPs´offices are organised. Most GPs work in groups of two ore more. Norway differs from may other countries, by having mostly only 2 professions in GP practices, namely GPs and health secretaries. (36) (37) However, there is a large diversity in how GPs organise and perform their work across the country. Some GP offices have dedicated nurses performing independent patient work and assisting the GPs. Differences in the patient populations, geographical conditions (including travel distances) and access to specialised care and other services have made local adaptations nessecary. In many municipalities, the GP service is co-located with the nursing station and mid-wife service. In more central areas with a higher population base, there are often larger medical centres, where GPs are organised in larger groups and are co-located with other services like psychologists, physiotherapists and private specialists like paediatricians, gynaechologists and dermatologists.

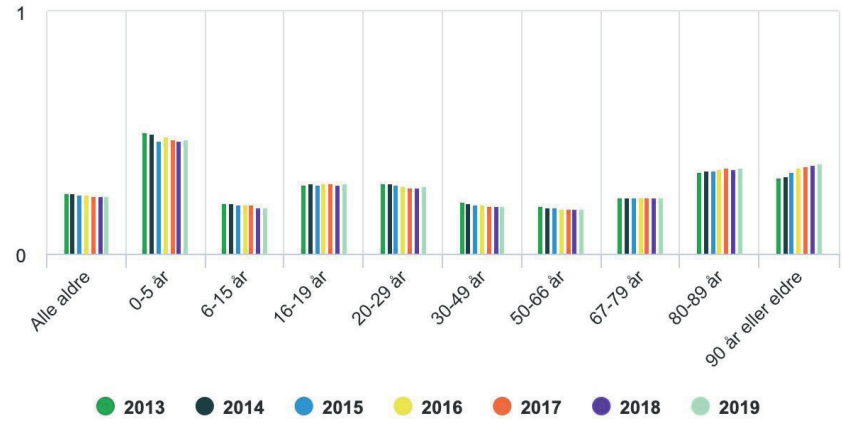
Even though the nature of General Practice is to provide generalist services, there is room for each GP to gain more specialised competence also in selected fields of their own choice. Often, personal interests an perceived local patient needs promote the engagement to build competence, gain more medical equipment and expand treatment offers. This is in line with the aims of the Coordination reform, as increased/adapted competence locally may prevent the need for specialised services. (6) It is also reflected in the many different offers for additional education for GPs. This development results in a large diversity of offered investigations and treatments across the GPs offices. For example, there are GPs performing diagnostic ultrasonography, group exercise and cognitive therapy, while others are trained in treating chronic wounds. There are also large variations in the performance of minor surgery and gynaechological exams.

2.2.2 Out-of-hours GP services

The out-of-hours services is intended to secure the population access to primary care for acute health care out of office hours. The municipalities' responsibility to provide out-of-hours services is described in the Regulation on organisation of emergency services (38). This requires that at least one physician is available to meet the emergency medical needs of the population also out of normal working hours. The out-of-hours services serve an essential role in the emergency medical services chain (prehospital services), with the Emergency medical communications centres and the ambulance services.

The GPs' contract with the municipality regulates their participation in the out-of-hours services. (26) For GPs, participating in the out-of-hours services in the municipality they work in is mandatory. Age over 60 years, pregnancy, nursing children under one year, and physical impairments or other weighty reasons, particularly for GPs over 55 years of age, allow the GPs to opt out. A study from 2018 showed that 2/3 of the GPs participated in out-of-hours work during one week in January 2018 (39). In 2019, about 60% of out-of-hours claims were made physicians registered in the regular general practitioner register, and about half of these were specialists in General Practice. (40) Other physicians staffing the out-of-hours services are mainly interns, locums, hospital specialists or physicians currently working in academia. (41) 35 % of the claims were made by female physicians, and about 25 % were made by physicians aged 50 years and older.

In 2019, there were about 2.15 million claims from patient contacts with the out-of-hours services. (40) Of these 62% were consultations, and 32% telephone contacts. The share of the population using the out-of-hours services is highest in small children, adolescents and older people. (**Figure 1**)



Kilde: Allmennlegetjenesten, Statistisk sentralbyrå.

Figure 1. Mean number of out-of-hours consultations per year (2013-2019), according to age group. (Statistics Norway)

In 2019, about 30 % of children aged 0-5 years had at least one out-of-hours contact. The timing of the use reflects the interplay with the regular GP service, as most out-of-hours contacts are made when regular GPs are unavailable. About 40 % of contacts are made during weekends, and the busiest period during the years are public holidays like Easter and Christmas. (**Figure 2**)

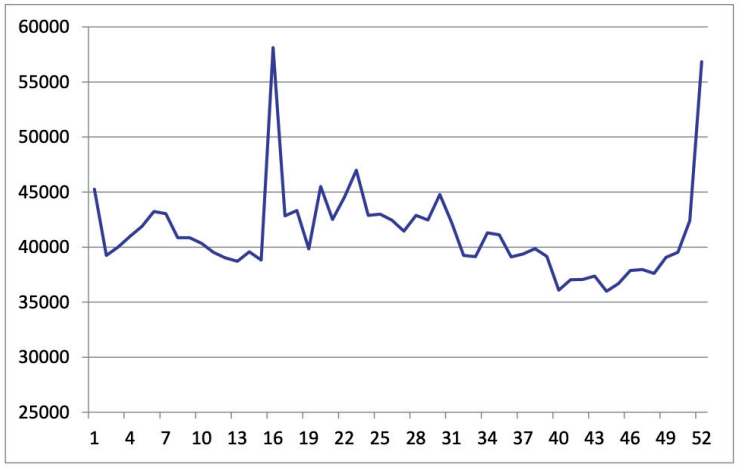


Figure 2. Number of out-of-hours contacts per week 1-52 during 2019. (Yearly statistics for out-of-hours services)

There is a large diversity in the organisation of out-of-hours services in different countries. (42) The GP cooperative model provided in Norway is also the most common model in Europe. However, compared to other countries, one of the main features in the Norwegian system is that patients do not have direct access to emergency departments, resulting in higher use of the out-of-hours services. In Norway, as in Europe, there is a trend towards upscaling and centralisation of the out-of-hours services. (40, 42)

Diversity in the out-of-hours service

There is a large diversity in how the municipalities organise their out-of-hours services in the different regions, and there are continuous adaptions to the changing needs and available resources. In 2018 there were 177 out-of-hours services, by which 102 were cooperations by several municipalities, and 75 were by single municipalities. Further, there were 97 local (municipal) emergency communication centres in Norway. (40)

The practical organisation ranges from only one physician being available on-call for requests by telephone, contacts at the doctor´s office, or home visits, to dedicated out-of-hours stations continuously staffed with specialised nurses and medical secretaries,

and several physicians working side by side. Further, there is a substantial diversity across the out-of-hours stations in both available medical equipment and facilities for clinical investigations, observation over time and treatment. These differences are mostly driven by the demographical conditions and needs in the specific areas. Scarcely populated areas may for instance have low staffing but highly equipped facilities, as there are a low number of patient contacts, but challenging geographical circumstances with long and potentially problematic travel distances to reach specialised care. In some regions, the municipalities have cooperated on organising community hospitals to secure sufficient treatments of patients when in need. When there is doubt about the need for specialised care relative to the efforts needed to transport the patient to the nearest hospital, available local facilities for observation and treatment may serve as an option.

This organisational diversity makes comparing the services in different areas challenging. From 2006 there has been yearly reporting of statistics on the provision and use of, and trends for the out-of-hours services, based on registry data from the Control and Payment of Health Reimbursement register (please see section 4.3.1). These reports have provided a valuable overview of the organisation and development of the service, and the data serves as a base for research in primary care. Unfortunately, there is no available national registry data for neither the emergency medical communications centres nor the ambulance services. (13)

There is also a substantial diversity in the use of the out-of-hours services. This will be further described in section 2.6 Variation in health service use.

2.3 Specialised health services

The provision and organisation of specialised health services is regulated by the Specialist health services act from 2001. (8) Specialised health services comprise inpatient care in hospitals and outpatient care provided by specialist outpatient clinics located at hospitals and other locations. Most Norwegian hospitals are publicly funded

and owned by the state; however, a small number of hospitals are privately owned. Norwegian legislation ensures free hospital choice. (8)

Specialised health services stand for about 90 % of the Norway´s health care expenditure, with a cost of about 151 billion NOK in 2019, increased by about 20 per cent from 2008. (5) In the latest years, both costs and activity levels in somatic hospitals and regional health authorities have steadily increased. From 2015 to 2019, the costs increased slightly more than the activity, resulting in a slightly lower estimated productivity level, with a reduction of 0.6 %. (5) Improving resource allocation by promoting a shift from inpatient to outpatient care is reflected by a decrease in the number of hospital beds, of 5.9 % from 2015 to 2019.

In 2019, about 2,5 million patients (46% of the population) used specialised health services. This share has been relatively stable over the previous five years; however, the share of elderly patients (above 67 years) increased in the period. (5) As expected, inpatient treatment decreased, and outpatient treatment increased over the five years from 2015 to 2019. In 2019, 521 282 unique patients were registered with at least one hospital admission, constituting about 10 per cent of the population (decreasing from 2016). The number of planned days in hospital were also reduced by almost 5% (10 000 days) from 2015 to 2019. Unplanned hospital stays were first reduced, then accordingly increased in the same period, and constituted 62% of all days in hospital (about 1 percentage point higher in 2019 compared to 2015).

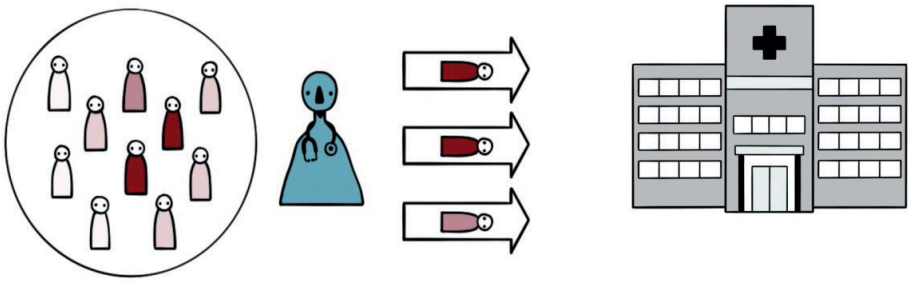


Figure 3. Gatekeeping

2.4 Gatekeeping

In Norway, as well as in many other developed countries, the organisation of the health services is based on primary care physicians as gatekeepers. (Figure 3) This implies that patients as a main rule, cannot access specialised health care without a referral from a primary care physician or hospital physician. Exceptions from this rule are made in emergency situations when patients are brought to hospital by ambulance, and for specific medical situations like child deliveries and sexually transmitted diseases.

In 2014, 64% of the 497 845 patients admitted to somatic hospital had been in contact with a GP, 28% with a regular GP, and 36% out-of-hours. (43) The remaining were either brought in directly by the ambulance service or had been referred by a physician not claiming public reimbursement.

2.4.1 Referrals

Gatekeeping founds upon the concept of referral. Despite all the legal actions a referral prompts, there is a lack of a formal definition in the international literature. However, for all means, a referral can be defined as a formal request for another health professional or institution to investigate or treat the patient´s health problem. This request can be based on a wish for treatment or special investigation that the physician cannot provide, an advice on diagnosis or management, a specialist declaration, or a

wish to share the medical responsibility. (44) For emergency referrals, the request often also involves assessment of the need for hospital admission.

In Norway, a national guideline defines what information a referral to the specialised service should contain, (Henvisningsveileder) (9) Further, there is a corresponding national guideline based on several legal documents (8, 26, 45) for use in the specialised services when prioritizing the referrals (46) This documents states the priority that should be given medical conditions of both urgent and non-urgent nature.

Defined by the Prioritizing Act (45), there are several criteria for entitling a patient the right to specialised care:

- A need for specialised health care
- An expected benefit from the health care
- There must be a reasonable cost-benefit ratio
- The health care is within the scope of the services the specialised services are obliged to provide and finance.

Based on these criteria, the referral is either approved or rejected. Importantly, there is no register of referrals in Norway. However, an inquiry by the The Office of the Auditor General suggested that from 2012 to 2016, the numbers of approved referrals registered in specialised services were relatively stable with more than 1.2 million per year. How many referrals were rejected in this period is unfortunately not known.

Elective and acute referrals

Referrals for non-urgent conditions are often referred to as elective referrals in the literature, while referrals for urgent conditions are referred to as acute, emergency or unplanned referrals. In addition to the obvious difference in the urgency, there are some essential distinctions between these situations.

In cases of elective referral, the assessment of the patients´ needs and rights for specialised care are mainly based on information in the referral letter sent to the specialist institution. The quality of these referral letters has been studied and debated.

(47, 48) If the patient is entitled to specialised services, this leads to an elective or planned appointment, within a certain time limit, based on the need. (45) In 2003, about two million elective referrals were issued in the Norwegian Health services, contributing with substantial assessment work for physicians. (49) About 90 per cent of these were assessed to entitle the patient to specialised health services, based on the referral letter, thus 10 per cent were rejected. In cases of rejection, the referrer is left with the responsibility for adequate follow-up of the patients.

However, in most cases of acute referrals, there is no assessment of the referral letter prior to the arrival of the patient in the specialised service. Thus, basically all acute referrals lead to unplanned contact with the specialised service, and require a new patient assessment by a physician, regardless of the quality or appropriateness of the referral. The specialist physician determines both the needs and rights for specialiced treatment and forms a plan for further care. For urgent cases, hospitals encourage the referring physicians to contact the hospital (mainly by phone) prior to sending the patient, both to allow preparations for the arrival, and to discuss the need for admission or if there are other potential options for the patient. (50, 51) There is currently no registration of how many of these contacts that leads to other solutions than sending the patient to the hospital. However, these contacts are seen as a target for reducing pressure in specialised services. Consequently, initiatives are taken to strengthen and systemize such requests/contacts, by providing primary care physicians with hot-line numbers to specialists, lowering the response-time, and securing the availability of experienced physicians to answer and give advice to the callers.

In this thesis I will emphasise the referral for unplanned health services.

2.4.2 Aspects of gatekeeping

Gatekeeping is a key feature of health service systems in many western countries, like Ireland, Italy, Netherlands, Portugal, Spain, Australia, New Zealand, Canada and the UK. There are some obvious pros and cons, which have repeatedly been subject for both debate and research. (52-54) A result of the selection process, the number of patients in the emergency departments are lower, but the proportion of these who are admitted to

the hospital is relatively higher, than in countries without gatekeeping. By limiting the inflow of patients, gatekeeping is shown to reduce both costs and pressure, thus improving access by limiting waiting time. (53) (54, 55) Also, patients may be protected from the potential harms of overdiagnostics and treatment. (56) On the other hand, possibly delayed entry to specialised care can in turn lead to deterioration of both chronic and acute health conditions and delayed detection of cancer, resulting in both rising costs and compromised patient safety and satisfaction. (54) This fine balance between referring too many, giving rising costs and pressure, and too few, compromising patient safety and satisfaction has received much attention. The utility of gatekeeping is, however, not only a question of referring too many or too few, but referring the correct patients, similar to the concept of diagnostic accuracy. This will be elaborated further in section 2.6.1 Variation in referral practices.

2.5 Assessing quality in the health services

In line with the description provided above, the health services constitute a large system with complex dynamics, where many factors are closely, but not always visibly, connected. Increased pressure in primary care, with for example longer waiting times to see a GP, may lead to increased pressure in the out-of-hours services, providing less time for assessment, or even deterioration of patients' conditions, leading to higher admission rates. Also, lack of available beds in nursing homes or resources for home care, causing patients to stay longer in hospitals despite being defined as dischargeable, may cause delayed treatment for patients waiting to be admitted.

To assess how changes - intended or not - made in one part of the service affects the other parts, there is a need for counting and measuring not only costs, but also outcome data like numbers of patients successfully treated, admission rates, or survival rates. Seen in context of changes in the population's needs, and in the available resources, such information is crucial for planning and dimensioning the health services.

Approaching large health care systems with the aim of quality improvement, comparing measurable outcomes such as admission rates or survival rates, may seem appealing,

and are widely used. Comparisons of outcome and process measures are often also the base for health services research. However, for drawing causal inference, and adjusting practice according to such comparisons, there is reason to be cautious, as this may not always be as simple and straightforwards as it appears. (57) It is argued that the complexity of health services dynamics, makes it hard or even unlikely to find valid models of comparison between providers, even after adjusting for all known measurable confounders. (58-60). However, in controlled situations, like in randomized controlled trials, health outcomes are well suited as quality or performance indicators.

An example from Norwegian health policy is the use of survival rates as a quality indicator, with a defined goal of increasing survival for specific diagnoses with relatively defined diagnostic critertia. In the whitepaper 'Norwegian Health and Hospital plan 2020-2023', there is a defined objective to increase the 30-day survival rates for hip fractures, by two per cent over five years. (61) In this case the survival rates will most likely be a fair indicator of how these hip fracture patients are treated in the health services, and may serves as an indicator of specialiced service performance. However, these rates are not unaffected by possible changes in the incidence of hip fracture, or changes in the age composition or comorbidities of the patients. Changes in both preventive measures and follow-up in other parts of the health service may also have implications for the measured survival rates, and thus give the specialised services undeserved credibility or discredibility. More is the objective of increasing the "total 30-day survival rate after hospital admission" by three per cent during the same 5 year period. This measure will have a much higher sucseptibility for impact from other factors not reflecting hospital quality, like an influenza more serious than usual, or the current pandemic. Thus, I argue that adapting the organisation of the hospitals based on general survival rates, may be harmful. Such quality indicators may, theoretically, also serve as incentivces for lowering the threshold for admitting patients with only minor health care needs, as this will automatically improve the survival rates.

With these potential pitfalls in mind, we have tried to find approaches potentially less susceptible to confounding factors. This will be discussed further in the next sections. In our project we have used practice variation both in the means of different provision

of health services, and further as a measure of different use and patient outcomes. There is a wide literature on practice variation in health services, but for this thesis, I will only present literature regarding variations in referrals and in health services use.

2.6 Variation in health services

Variation in the delivery, use and outcome of health services is the main foundation for health services research. (62) By observing variations in health services delivery and comparing outcomes, Dr Ignaz Semmelweiss and Florence Nightingale both contributed with crucial progress in the understanding and organisation of health services. (63, 64) This has been followed by more structured approaches to investigate dynamics in health services. (65) More recently, John Wennberg has contributed substantially to shedding light on variations in modern health services and understanding supply-sensitive and preference-sensitive care variation. He has further emphasized the need to monitor variation in the delivery and outcome in the services and that as long as higher use of resources does not result in higher improved outcomes and quality of life for the population, is it of high importance to reduce the variation. (66)

With a strong emphasis on equitable access to health care, striking differences in health services use and outcome both within and across services and nations have been regarded as an indicator of some form of non-optimized practice, suggesting that both too little and too much may be harmful. Consequently, reducing variation has been appointed a target for quality improvement. However, distinguishing between *observed* variation and *unwarranted* variation is essential. Differences in geography, demography, morbidity, and quality in other services often influence variations in health services use and patient outcome. Thus, variation is considered unwarranted when it is not caused by such differences in health care needs. (66)

Nevertheless, identifying this unwarranted variation by observing and monitoring complex health systems can be quite difficult. The likeliness of keeping control of all possible mechanisms of variation without using standardized situations like in randomized controlled trials has been questioned. (67, 68) Consequently, using

observational data requires research designs addressing these challenges of possible confounding.

Moreover, implicit in the mission of reducing variation lies the question of identifying the correct level of care. Based on the different needs in the population and the functionality of other available health services, the optimal level of care may vary substantially between different groups. A necessary adaption to these needs and available resources may be an important mechanism contributing to the variation. An example of this may be comparing varying admission rates for pneumonia for elderly patients between different municipalities. After adjusting for age and co-morbidities of the patients and travel distance to hospital, differences in the GPs judgements remain the point for intervention to reduce the variation. However, if no additional knowledge on the other optional services and alternatives to admission is available, we might end up with the wrong target, as the variation in rates may also depend upon, e.g., the availability of municipal acute wards, the functionality of homecare, and the opportunity for the GPs to make home-visits. If patients in the municipalities with high admission rates have poorer alternatives to admission, trying to reduce this variation may cause more harm than good, despite ideal intentions. Thus, I will suggest a careful approach to using variation as performance indicators also in health services research. (69)

In Norway, variations in health services use have become increasingly imperative on the political agenda. In 2004 the “Centre for clinical documentation and evaluation” (SKDE) was established, to illuminate the geographical inequalities in specialised services, contributing to equitable and high-quality health services regardless of where the patients live. (70) Since then, the centre has published yearly reports and research on variation in Norway’s health services use and outcome. From 2015 the centre was asked to provide a national health care atlas to provide updated information on health services use and variations, resulting in the webpage www.helseatlas.no.

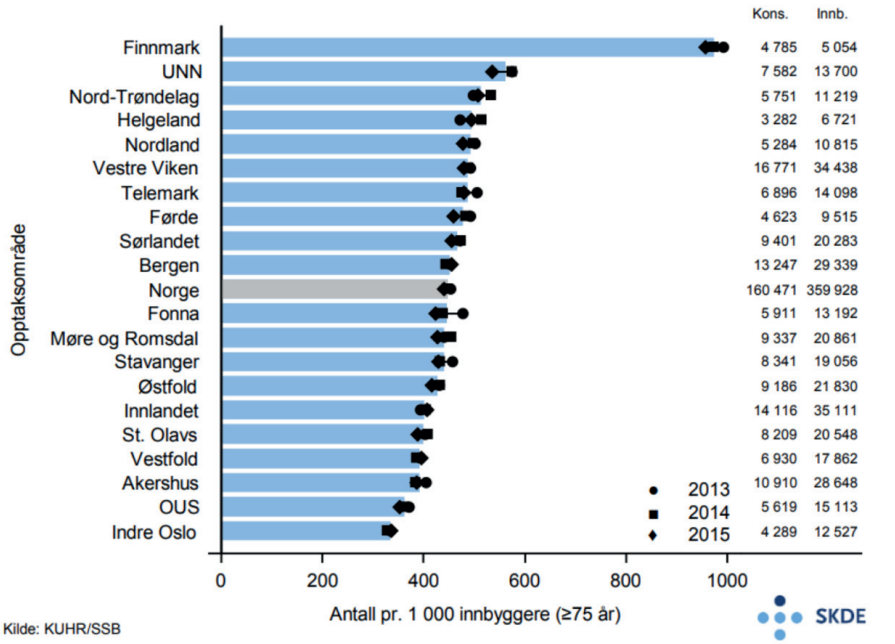


Figure 4 Out-of-hours consultations per 1000 citizens, according to hospital uptake area, adjusted for age and sex, per year 2013-2015. (www.helseatlas.no)

Data from the health care atlas shows substantial variation in both the provision of investigations and treatment offers, and surgical procedures, outpatient treatment, and hospital admissions across the country. As an example, Figure 4 show variations in the use of out-of-hours services related to geography, with a substantially higher use in the uptake area for Finnmark hospital. Such findings are used as a base for targeting and investigating unwarranted variation and implementing measures to reduce such variation.

Furthermore, identifying variation in health care delivery and defined outcomes is one side of the coin; however, investigating the extended consequences resulting from this variation is also highly relevant. In the existing literature, the long-term impact of variation has received varying attention. A suitable example of this is differences in referral rates. As a wide range of studies has found substantial variations in referral rates

between physicians, how these variations impact the patient outcome or further health care use is more scarcely studied.

2.6.1 Variation in referral practices

Considering the pivotal impact of the gatekeeper function for access to and use of spesialiced services, variation in referral rates have been of high interest over decades, both nationally and internationally. Researchers have approached this matter from many angles, with varying methods; observational analyses, questionnaires, vignette studies, and qualitative interviews, to mention some. This broad approach fits well to this complex matter. The referral decision is made in widely different circumstances and with so many needs and aspects to consider that it may be hard to find suitable comparable situations. Nevertheless, making these decisions is a crucial part of the physician practice, with substantial consequences for both patient, system and physician, so it should be optimised.

Referral rates

Referral rates have increased substantially over the latest decades, both in Norway and many other countries. (71-73) For both elective and acute referrals, observational studies have shown substantial differences in physician referral rates. (12, 74) (75).

Consequently, what contributes to these observed differences has been subject to much research. (10, 11, 76-78). Varying needs in the population like age, gender, and socioeconomic deprivation are suggested to contribute substantially (11, 59, 77). Further organisational factors, such as practice characteristics, access to care, and distances to hospital are also proposed to impact referral rates. (79) However, a large share of the variation remains unexplained, (59) and has been suggested to be influenced by physicians' characteristics and decision making. (77) Studies have found that physicians' sex, age, experience, and specialist status impact referral rates; female physicians, younger, more inexperienced, or non-specialist physicians are associated with higher referral rates. (10, 74, 75, 80-82) In the following, I will mainly present

literature relevant for acute referrals and highlight the suggested impact from physician characteristics.

Acute referrals

The situation of an acute referral is often characterised by a lack of time because of the acuity or severity of the patient´s health condition and may also be perceived as having a higher consequence for the patient. Further, many acute referrals are made in out-of-hours settings, which differs from the conditions for regular GPs working daytime. The out-of-hours physician often has scarce or no knowledge about the patient, and limited time, resources, and access to medical records or conferring with a specialist. Also, in contrast to a regular GP, the out-of-hours physician gets little or no feedback on the referral outcome. (76, 83) Nevertheless, out-of-hours physicians express worries about negative consequences like lawsuits or media attention in case of adverse outcomes. (76)

Out-of-hours services provide a good context for studying variations in referrals, as different GPs are serving the same patient population, in contrast to GPs serving their "own" selected patient list during regular working hours. In the out-of-hours context, impacts from differing patient populations and organisational factors are likely smaller (though still present). Also, as the morbidity of the out-of-hours patient population is higher, the differences in referral rates are likely accentuated. A British study from 2007 found a five-fold variation between out-of-hours physicians serving the same out-of-hours population. (75) Consultations with a female GP, during night shifts (11 pm to 7 am), and home visits were associated with an increased rate of emergency referrals; however, these factors did not explain all the variation.

A qualitative interview follow-up study including GPs defined as high, medium and low referrers suggested GP characteristics such as level of professional confidence, tolerance of risk and uncertainty and views of potential alternatives to hospital admissions to be explaining factors in this association. (76) Other studies have followed this, exploring the association between physicians´ experience, risk management and

tolerance of uncertainty, (81, 84) suggesting that physicians with lower risk of tolerance also have lower threshold for referring.

Referral threshold

As a physician's referral *rates* may be highly affected by the morbidity in the patient population, the referral *threshold* may be a more consistent characteristic of the physician, and this may be hard to disentangle from other characteristics such as sex, age, experience, and specialisation. The referral threshold is believed to be related to personality traits like self-confidence and risk-aversion, but also to the approach to patient-centred practice and taking the needs of the next-of-kin into consideration. (78, 81).

An essential aspect of variation in referral threshold between physicians is for which patients this variation will apply. Most likely, a large amount of the patients in primary care will never be directly affected by the physicians´ varying thresholds, as they are clearly not in need of a referral, or on the contrary, obviously in need of a referral. However, for some patients, the referral threshold of the physician they meet will be decisive of their further care. It is also reasonable to believe that these patients will be most affected by general requests for primary care physicians to lower their referral rates.

Referral accuracy

As much as variations in referral rates have been investigated, the essence of referral practice is the *referral accuracy*. Hence, the physician's ability to identify the correct patients to refer, similar to the diagnostic accuracy of a test. (37) If two physicians have similar referral rates but differing referral accuracy, this would still be problematic, as the one with lower accuracy would incorrectly refer, or not refer, more patients. (85) Another closely related aspect is the physcians' confidence in their own assessment and referral decision. Both accuracy and confidence are suggested to improve with increasing experience; however, conflicting results are reported. (86, 87) Possible explanations listed are the lack of specific training on referral decisions, and feedback

on the outcome of the decisions. Enhancing referral accuracy is identified as a goal in the Norwegian health policy. Nevertheless, there is still room for suitable measures to optimise the framework for decisions and increase the learning outcome.

2.6.2 Medical overuse/overactivity

Closely related to variation in health services use are also the potential harms of too little or too much. As high-quality health services have been scarce, unequal access to health care has mainly been regarded as a problem for those who receive too little care. However, in the latest decades, medical overuse is increasingly recognized as a threat to public health in industrialized countries, and internationally, many initiatives aim to limit the potentially harmful effects of too much medicine. (88) In a systematic review from 2017, potential drivers for overdiagnosis is mapped and linked to potential solutions. (89) Many of the suggested drivers are closely connected to the general development of society and thus not easily changed. However, one approach is to ensure that the provided health services are beneficial for the population.

In Norway, 2019, the current Minister of Health Bent Høie, inspired by the English "Evidence-based interventions programme", (90) provided two assignments to the Reginal Health Agencies. The aims were to evaluate and ensure the evidence-based benefits of both implemented and new investigations, treatments, procedures and medical technology and resulted in the "Reconsiderations project" (91). Concurrently, the Norwegian Medical association started the Choosing Wisely Campaign, also inspired by an English campaign with the same name. (92) These initiatives aim to reduce the adverse effects of overuse of health services on both the patient and system level.

Although the primary attention is given to reducing unnecessary use of elective health services, overuse of acute and emergency services also imposes substantial risks for patients and costs and pressure on the health services. Hospitalization increases the risk of adverse events, (93) disabling (94), and incorrect medication after discharge (95) especially in the elderly and more frail population. Also, as the capacity limits in the

health services are reached or even exceeded, overuse of emergency services will have large consequences for other patients in need of the services.

2.7 Capacity limits, pressures, and workload

There is reason to believe that patient flow and treatment in primary and specialised health care are closely related. Most research situates mainly in either primary or specialised services, and the intersection between the two is more scarcely studied. However, evidence suggests that well-functioning primary care limits or prevents the use of specialised care. (20, 96-99) Also, more tasks and responsibilities are shifted over to primary care to relieve the specialised services of pressure. With ageing and increasing population, rising expectations and technological development, the demands on primary care are also reaching an all-time high, both in Norway and many other Western countries. (3) If this impacts primary care's ability to handle the patients, and threatens the gatekeeper function, this may lead to even higher pressure in specialised care. How this affects patient outcomes is not known.

Increasing pressures in general practice is suggested to threaten the sustainability of the workforce, and thus the primary care services, both nationally and internationally. (7, 35, 100-103) The crisis in primary care in England is probably one of the most well-documented (among the developed countries), with problems of recruitment and retainment, reports of difficulties in access to care, and declining patient satisfaction. (100) (103, 104) The clinical activity level has increased substantially over the latest years, arguably without accordingly increase in funding, and the main challenge reported by the GPs is that the total workload is too heavy. (100, 102, 105, 106) Concurrently, the pressure on the specialised services is rising year by year. (107)

Similar mechanisms of increasing workload in health services may be present in many countries. In addition to challenging recruitment and retention of the workforce, it is suggested to affect both service delivery and patient care through provider burnout. (108) *Workload* is defined by the Oxford dictionary simply as *"the amount of work to be done by someone or something."* However, quantifying workload in health care is

quite difficult, mainly because two patients are never alike. The number of consultations per day, for example, does not contain information about the amount of work or the emotional load imposed by each patient. The objective measure of reported working hours may be a better indicator of workload and can be held up against national standards for working hours. Nevertheless, also this measure lacks information about the how the working day has been, the work intensity, opportunities to take breaks, and the emotional restrain, affecting the overall perception of workload. Investigating the amount and further the determinants, effects and dynamics of workload would thus benefit from several different approaches. The potential impacts of increasing workload in primary care, such as inadequate patient care, can be temporarily compensated for/contained by other parts of the health services, such as increased use of specialised care through increased referrals. Such mechanisms may be hard to detect but might be elucidated by investigating the perceived causes and consequences of workload and individual coping strategies.

An increase in working hours for GPs has also been reported in Norway, both during regular working hours, with a substantial contribution from work without direct patient contact, and in the out-of-hours services. (32, 39, 109, 110) The latest years, increasing workload in General Practice has been brought to attention by several actors, both among the GPs themselves, by the media and later by politicians. (7, 33, 34, 111) Challenges with recruitment and retention of GPs are increasingly evident in Norway, and political actions have been taken to secure recruitment through offering permanent positions for GPs in primary care instead of self-employment practice. (112) Introducing primary health care teams with more General Practice nurses is also piloted to relieve the GPs of pressure. (113) However, if this will be sufficient is questionable. A rising number of patients are now affiliated with regular GP lists served by locums or without a defined GP, and there is reason to be concerned about how this impact both the direct patient care and the dynamics between primary and specialised health care. (7)

Also in the specialised services both nationally and internationally, there are steadily rising number of patient contacts, referrals and hospital admissions. (5, 107) Issues of

particular concern are the increasing reports of periodically overcrowded emergency departments, and capacity limits reached or even exceeded in the hospitals. (114-117)

Pressure in primary care may impact the gatekeeping role in the direction of higher referral rates, and possibly less accurate referrals. Pressure in specialised services may affect the gatekeeper role in the direction of lower referral rates through limited capacity in the services. How this interplay affects the accuracy of the referrals, i.e., selection of the correct patients to specialised services, is not known. Nevertheless, the effect of pressure on this interplay may have tremendous consequences on both the patients and the health services and warrants further research.

In this thesis I will look specifically into matters regarding the gatekeeping role of general practitioners and how this further affects the inflow of patients to hospitals, and in turn health services capacity. Paper I concentrates on causes and consequences of workload in general practice, Paper II on physician factors affecting gatekeeping, and Paper III on how a referral affects both the patient, the specialised services, and in turn the general practitioner service. (Figure 5) My hope is for this thesis to be a contribution to the essential work of preventing the fine dynamics in the health services from transforming into a vicious cycle of increasing pressure and limited capacity.

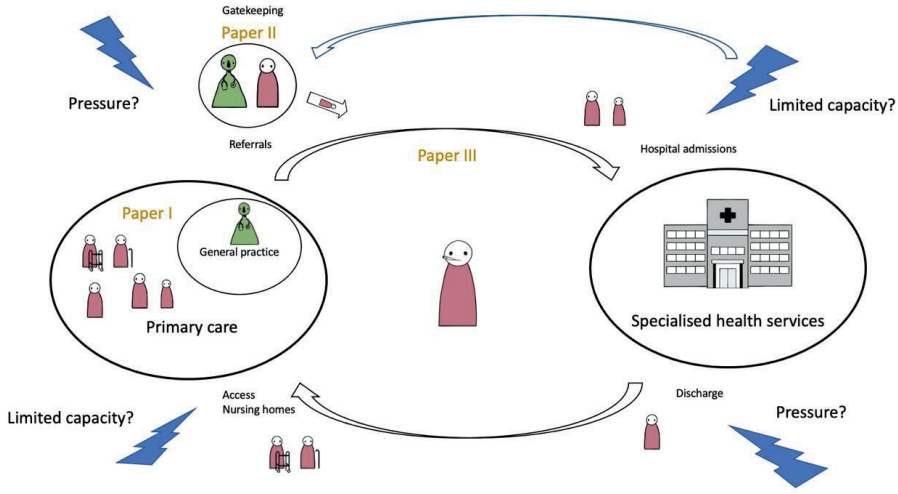


Figure 5. An illustration of how pressure and limited capacity may potentially impact the dynamics between primary care and specialised health services, and where our three papers are situated.

3 Aim of the thesis

The overall aim of thesis was to investigate and explore the dynamics of health services system, through studying causes and consequences of referrals for patient outcomes, safety, and further health service use. Physicians have differing thresholds for referrals and varying referral rates, and there is a lack of research on how these differences impact patient safety and health service use. There is also scarce if any knowledge on the effects of altering these thresholds. High pressure and workload may impact on the GPs´ and out-of-hours services´ gatekeeper function directly, or through affecting the composition of physicians staffing both General Practice and out-of-hours services. This may in turn induce more pressure on the specialised health service and possibly affect patient safety. Moreover, using observational data to assess the effect of being referred is difficult because of confounding by indication.

More specifically, the aims of the three papers were:

- To explore how general practitioners and their co-workers in Mid-Norway perceive the causes and consequences of variation in workload in General Practice. (Paper I)
- To investigate associations between GP characteristics, unplanned hospital admissions and patient safety for general practitioners working in an out-of-hours setting. (Paper II)
- To estimate the impact of altering the referral threshold from the out-of-hours services on further health care use and patient outcomes. (Paper III)

4 Materials and methods

This thesis includes three papers with two different methodological approaches: one qualitative and two quantitative studies. In Paper I we aimed to explore how GPs and coworkers in Norway tackle and perceive workload. We chose to collect data by combining qualitative individual interviews with focus groups, as this was the most feasible solution to best include the participants we wanted. We aimed to explore in which periods the GPs were affected by a high workload and further use this information to inform practice and to generate relevant hypotheses for further quantitative analyses in the project.

For the two quantitative studies, we had access to an extensive link of Norwegian register data. Based on a thorough assessment of the data and the research questions, we used different statistical methods to exploit the available data. In Paper II we carefully matched the data by defining comparable groups and used linear and logistic regression with within-group estimators to investigate the exposure-outcome associations. In Paper III, we used multivariable adjusted linear regression for description and instrumental variable analyses with within-group estimators with the aim of assessing causal associations. In the following, I will elaborate on the rationale for these choices and the details of the methods.

4.1 Project context

This thesis is a part of a larger project named *«Health care services under pressure – consequences for patient flows, efficiency and patient safety»*, (118) financed by the Norwegian Research Council. The aim of this project was to investigate different aspects of the health care services, and how situations with increased pressure both in primary and specialised health services can affect patient flow, efficiency, and patient safety. This was mainly a quantitative project based on the extensive link between several Norwegian health care and demographic registers, but also had a defined aim to conduct a qualitative exploration of the research field to generate good hypotheses for the quantitative research. In the original plan, the project period was defined to be

between the 1st of January 2017 and the 31st of December 2020. However, due to the pandemic and the arrival of my fourth daughter, the project period was prolonged.

When I started my PhD-period, I was immediately included in an interdisciplinary research group comprising competence in health services research, epidemiology, econometrics, and statistics affiliated with different institutions; Department of Public Health and Nursing (NTNU), Department of Economics (NTNU), St. Olavs hospital and University of Bristol. With group meetings every second week, research topics and questions were enlightened from different perspectives, and subprojects were presented and discussed.

4.2 Qualitative method

Many research questions cannot be answered by quantification alone and qualitative research methods may contribute with empirical evidence from different angles. (119) The main benefits of qualitative research are the possibilities to explore complex phenomena to generate new hypotheses and in-depth understanding of the field. Further, it also emphasizes the views and experiences of the participants. This makes qualitative methods well suited for studying how workload is perceived and tackled by health care providers. Qualitative research further provides valuable insight into the research field to inform practice and decision making. It may serve as both base for and as a supplement to the quantitative approach, as it did in our project.

4.2.1 Design Paper I

In Paper I, we aimed to explore how GPs' and their coworkers' (nurses/health secretaries) perceived and tackled their workload and their experiences and reflections regarding the explanations for, and consequences of periods of high workload in Norwegian general practice. We chose a qualitative approach as this is the best way to explore and provide rich descriptions of these complex phenomena. We decided to collect data through interviews with GPs and their coworkers in urban and rural municipalities in Mid-Norway.

The subject of interest, namely workload, is assumingly highly correlated with the potential participants' availabilities to and interests in participating in a study. If experiencing a particularly high workload, this could lead to declining the invitation, and such a selection process could negatively affect the variety and nuances of the results. Considering this, we strived for appearing most flexible in the invites to participate in the study, both regarding the time and place for the interviews. In some of the GP practices, not all the invited participants were available for interviewing simultaneously. Considering the possible disadvantages of losing these potential informants, and the advantages of the opportunity to explore and assess the dynamics and the statements given both in groups and individually, we decided to offer individual interviews in these cases. This resulted in using both focus groups and individual interviews for data collection, which also allowed for a more comprehensive approach to explore the experiences of workload. This will be elaborated on in section 4.2.3 Data collection.

4.2.2 Study setting and population Paper I

From September 2017 to January 2018, we visited 7 GPs practices, in 6 urban and rural municipalities in the two counties Trøndelag and Møre og Romsdal in Mid-Norway. The recruitment process was made by strategic sampling. Based on our knowledge about the varying GP practices in the county, we selected some criteria to secure a wide range of perspectives on the topic. Hence, we aimed to include GPs with varying sex, age, experience, practice size, managing style and geographical location. Potential participants received personal invitations by e-mail, including a PDF with a description of the project, the written forms of consent and the questionnaire for supplementary information on the participants. (Appendix II) Only four of the invitations were declined, and the reason given was lack of time. We performed seven focus groups, whereof three with GPs, three with coworkers, and one focus group with GPs and coworkers together. In addition, we performed four individual interviews with GPs. (Figure 6)

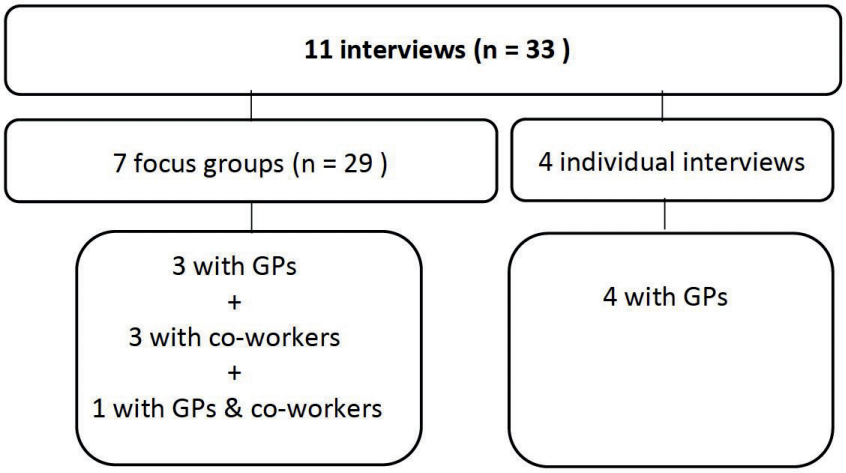


Figure 6. Overview of the data collection, Paper I.

This made a total of 23 GPs and 10 of their coworkers. Of the total 33 participants in this study, 21 were female, and 12 were male. Of the GPs, 11 were female, and 12 were male. Further details of the participants are presented in table 1 in Paper I. Table 1 shows the participant characteristics, provided on a group level.

Table 1. Participant characteristics, provided on a group level for the focus groups and individual interviews.

	Location		Occupation		Gender		Participants
Focus group n= 7	Urban	Rural	GP	Coworker	Male	Female	
	x		7	0	5	2	7
	x		3	3	2	4	6
	x		0	2	0	2	2
		x	0	2	0	2	2
		x	6	0	2	4	6
		x	0	3	0	3	3
		x	3	0	2	1	3
Individual interview n= 4	2	2	4	0	1	3	4
Total	5	6	23	10	12	21	33

The interviews took place at the different GP practices, during or after working hours, based on the participants' wishes. All participants signed a written consent to participate. They also provided supplementary information about their gender, age (in categories), and characteristics about their workplace. GPs were asked about the number of years working as a regular GP, their specialist status, number of patients per list, days of the week at the GP practice, time scheduled per patient, and special municipal responsibilities. Coworkers were asked about the number of years as a medical secretary/nurse, education, number of health care personnel in the practice, average working hours per week, average patient contacts per day, and if there were other features characterising their GP practice. (Appendix II) In order to maintain the anonymity/avoid the possibility of participant identification, some of these details were omitted both for the paper and for this thesis. The participants received a small gift of gratitude after the interviews.

4.2.3 Data collection

There are several different techniques to collect qualitative data, (e.g., observation, focus groups, in depth-interviewing), all with the common feature of observing and/or interacting with the participants in the study. (120) Qualitative interviewing relies on the researcher's skills and views to produce data in the relation between the interviewer and the study participants/interviewees. (120) Individual interviewing is best suited for studying individual views and experiences and can also be used for sensitive topics where a group setting would lead to potential self-censoring. In individual interviews the interviewee can trust full confidentiality. In contrast, focus groups are recommended for exploring experiences shared by several participants, and where the group dynamics can encourage different stories and perspectives of the same phenomena. Here, confidentiality can be encouraged but not fully obtained.

For the focus groups, we wanted the benefits of the participants' close professional relationship which provide a setting of confidentiality and trust. Further, we appreciated the potential differences between GP practices and chose not to combine participants from different GP practices. Due to this, in addition to some urgent cancellations from participants, we ended up with fewer participants than we aimed for in some of the focus groups and had to convert to individual interviews in four cases. (Figure 2) However, as these interview settings contributed to our understanding of the topics and themes and did not deviate substantially from the focus groups, we chose to include them in the study. This will be further elaborated in the discussion section.

For the focus groups we chose to perform three different sets of participants:

1) GPs only
2) Coworkers only
3) GPs and coworkers together.

In the groups with GPs or coworkers only, we wanted to exploit the dynamics of participants with different views and experiences from more or less the same setting or point of view, like being a GP or coworker. This allowed the participants to describe the phenomena of interest and the others to fill in with their nuances. Combining both the GPs and coworkers in one group, we hoped to allow for different perspectives and points of view on the same phenomena.

For the work with this qualitative study, the co-authors formed a smaller qualitative research group, consisting of four physicians, one social scientist and one economist, which contributed to reducing reduce single researcher bias. Also, to allow for all the co-authors to gain proximity and better understanding of the data, and provide better value of the analyses, five of the co-authors also participated in the data collection, at least two in every interview session. Taking part in the data collection allows the researchers better conditions for experiencing the data and can contribute to providing ideas in the analytical process. (120) We altered the position as assistant and main interviewer/moderator for the different interview sessions. We also discussed and gave each other feedback on the moderator role throughout the data collection period.

There are different degrees of structuring a research interview; however, the most common is semi-structured interviewing, where the interviewer(s) follow an interview topic guide, allowing the participants to elaborate freely on topics that they are concerned with. We used a semi-structured interview guide, that was previously pilot-tested by academic GPs at our department. (Appendix I) The interview guide was adjusted several times throughout the interview period. This choice based on the observations of the participants consequently seeming to lead the discussion over from entailing "particular periods of high workload" to describing their experience of an "ever-increasing high workload". They proceeded by discussing the mechanisms and consequences of this perceived trend, rather than identifying what caused and came out of periodically increased workload. By adjusting the interview topic guide, we allowed the participants in the later focus groups and individual interviews to elaborate on the increasing workload, and their thoughts about how this affected their daily life and perspectives of the future. This observation and adjustment also informed the later identification of themes, analyses process and aims.

Focus groups

The group sessions started with the moderator informing the participants about the purpose of the study, informed consent, and the ability to withdraw at any point. The participants were also informed about the audio recording and about secure storage of data. Introducing one topic at the time from the semi-structured interview guide, the moderator allowed the participants to interpret the questions themselves and encouraged them to speak freely and share their experiences and views openly in the group. Our impression from all the focus groups was that the topic workload was very engaging for the participants, and that they felt comfortable in the setting, as they all took part in the conversation, also when they disagreed with the expressed views. The participants' statements were followed up, exploring if the same experiences were shared by other participants.

Individual interviews

Using the same semi-structured interview guide, we performed four individual interviews with GPs. These sessions also started by informing about the purpose, consent, ability to withdraw, audio-recording and data storage. Further the interviewees were given room to talk about the themes of interest and follow up questions encouraging them to elaborate on the topics.

Both individual interviews and focus groups lasted approximately 60 minutes. Field notes were made during and immediately after the interviews. At the end of each session, the participants were asked to fill out the questionnaires with complementary information, as described above.

4.2.4 Qualitative analysis

To analyse the data we used Systematic Text Condensation. (121) This is a four-step descriptive analytical process aiming to explore and describe the participants' experience as they express them, inspired by Giorgi's descriptive phenomenology. It includes reduction of data, and a shift between de-contextualisation and re-contextualisation. (121) It is regarded as well-suited for novice researchers while still ensuring a responsible level of methodological quality. We followed the analytical steps as outlined by Malterud.

Step 1) "Total impression – from chaos to themes"

All authors read the data material separately to obtain an overall impression of the data. After conducting, transcribing, and reading through the three first interviews, we gathered for a group meeting, where the process and progress so far were discussed, and preliminary themes were identified. It was at this point we became aware that the participants and interviewees consequently chose to lead the discussion over from periods of high workload to a perception of a constant increase in workload. In the following, we conducted and transcribed the remaining interviews before listening to and reading all the material.

Step 2) "Identifying and sorting meaning units – from themes to codes"

In the next step, the themes and content were discussed in our researcher group, and we agreed on a coding frame with five main categories; Causes of workload, consequences of workload, strategies for dealing with workload (both performed and suggested), reasons for staying/resilience, and interaction with specialized health care services. (Despite not being explicitly reflected in the titles of the main categories, we all agreed to the understanding of "workload" as "*increasing* workload", and not "*periodically high* workload" as we first had intentions of exploring.) Further, all passages were coded in these categories and sub coded into new categories. Codes and topics were thoroughly discussed in the researcher group and adjusted through the analysis.

Step 3) "Condensation – from code to meaning"

The following step included condensating and abstracting the meaning unit within each of the coded groups. Meaning unites were condensated into "artificial citations" written from the first-person perspective, which is the essential part of systematic text condensation. For this process, we shared the main responsibility for the five main categories between the researchers and had regular meetings and discussions during the process.

Step 4) "Synthesizing – from condensation to descriptions and concepts"

In the final step, the condensations were synthesised into major topics and sub-topics that reflected the participants and interviewees´ experiences of causes and consequences of their workload. This process led directly over to writing the paper. We made a final selection of illustrative quotes suggested during the process, and these were translated from Norwegian to English to fit the paper.

4.3 Quantitative methods

4.3.1 Paper II & III

Paper II and III are based on an extensive data set linking Norwegian public health services registers to registers of demographic information about the patients and

characteristics of the physicians. This data set allowed us to link physicians and patients' characteristics to health service use and patient outcomes. For our studies, we extracted all contacts with physicians claiming public reimbursement in primary care and all contacts with specialised health services in Norway from 2008-2016, except contacts with and admissions to psychiatric specialists and hospitals. For both papers, the linkage was based on the unique 11-digit personal identification number provided in all of the registers. Linkage of physician characteristics was based on the unique physician identification number (Helsepersonellnummer, HPR). All data were de-identified upon provision.

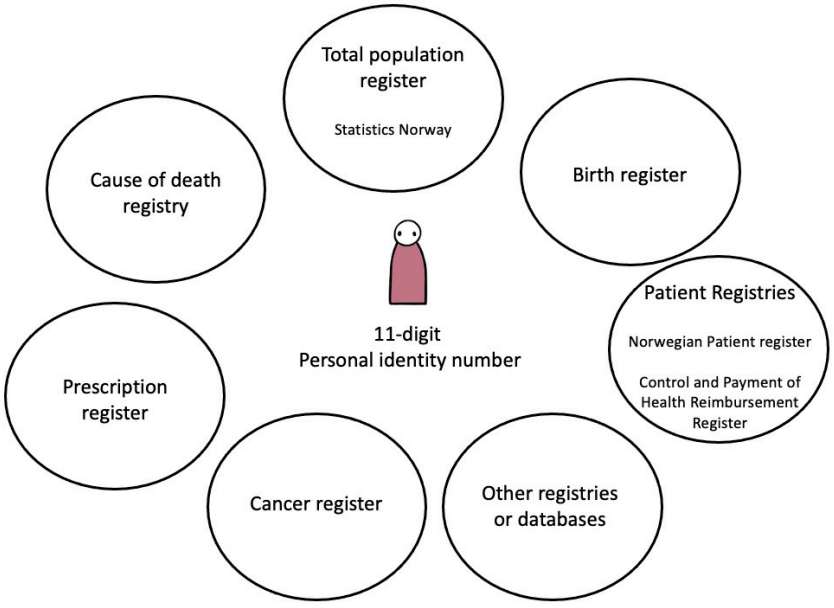


Figure 7. An overview of health services and demographic registers in Norway, Figure by Laugesen, 2021, (modified).

Registry data

Registry data from population-based and health care registers are becoming increasingly important and appreciated in research. (15) In Norway, all residents are assigned a unique 11-digit personal identity number upon birth or immigration, following them through their whole lives. This system, initially introduced in 1964 for control and monitoring of tax payments, has increasingly been used for administrative purposes by authorities, health services systems, schools, and banks, ect. The extensive use of the personal identity number is considered a gold mine for registry-based research, as it enables accurate individual-level linkage of different registries. (Figure 7) In the following, I will provide a short description of each of the registers we have used for our quantitative studies, as shown in Figure 8.

Control and Payment of Health Reimbursement Register (KUHR)

All patient contacts with Norwegian publicly funded general practitioner services, both during regular working hours and out-of-hours, generate a unique claim for reimbursement. These claims are submitted from the treating physician or service to the Norwegian Health Economics Administration and are registered in the Control and Payment of Health Reimbursement Register (KUHR). (122) The claims include the patient's personal identity number, time, patient diagnoses (ICPC-2), the unique physician identification number, and the type of contact, (e.g., telephone, consultation, home-visits, including a specific code for out-of-hours work), as well as codes for procedures. This register serves as the base for both Papers II and III.

The Norwegian Patient Register (NPR)

The Norwegian Patient Register (123) is provided by The Norwegian Directorate of Health and serves as a base for monitoring patient activity in the specialised health services and for funding. From 2008 the patient's personal identity number was included, thus allowing linkage to other registers. Some of the main variables provided by this register include the name of the hospital/institution, time and date for contacts and inpatient stays, level of care, urgency grading, codes for diagnoses (ICD-10) and procedures (NCPC-2). It also includes diagnosis-related groups (DRG) points for generating activity based-funding.

The Norwegian Cause of Death Registry (DAR)

The Norwegian Cause of Death Registry (124) is provided by The Norwegian Institute of Public Health, with the aim of surveilling the causes of death in the Norwegian population. This register provides the variables personal identity number, date of death, and cause of death.

Statistics Norway (SSB)

Statistics Norway (125) provides a total population register, including variables like date of birth, emigration and death, municipality of residence, immigration background, and educational level for the entire population.

The Norwegian General Practitioner Register (FLR)

The Norwegian General Practitioner Register is provided by The Norwegian Health Economics Administration (Helfo), (126) and contains data on physicians working as regular general practitioners. Initially, only the regular general practitioners were included; however, locums have also been included over the years. From 2016 all physicians claiming reimbursement from the Control and Payment of Health Reimbursement register are registered for work in the general practitioner services. (40) The register provides variables such as the physician identity number, the physicians' date of birth, gender, specialist status, and patient list affiliation (with patient personal identity number).

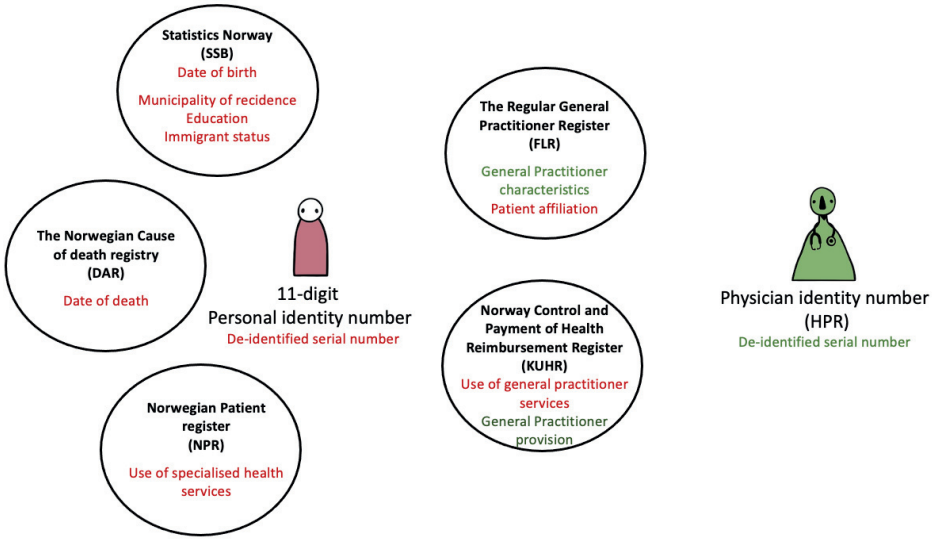


Figure 8. Overview of the national registers we have used, and the data extracted from the respective registers. Patient data in red, physician data in green.

Research methods - Using observational data for causal inference

For researchers, a well-known aim is to help the world move one step forwards by producing reliable facts about causal inference. However, overcoming the challenges of confounding factors is not easy. A confounding factor or *confounder* is a variable known to be a common cause of both the exposure and the outcome of interest. Studying the exposure-outcome association without handling such confounders may cause spurious associations between the exposure and the outcome, and thus, invalidate the results. Examples of confounders relevant for this thesis, are patient factors, like age or severity of a health condition, affecting both the patient's chance of being referred to a hospital, and using health services in the future. These are observable variables; however, not always easy to quantify. If measurable in a way we believe to be valid, it is possible to try adjusting for these confounders in the statistical analyses, to reduce the risk of biased results. However, there are often unobservable confounders, e.g., the patient's next-of-kin's convincing speech about why this patient needs a hospital referral. This next-of-kin may also influence the care this patient receives in the future.

Unmeasurable or unobservable confounders are infamous for their limiting effects on the use of observational data for causal inference.

The randomized controlled trial is still the gold standard for studying causality, with the central concept of random assignment to treatment in two (or more) groups and comparing the outcome distribution in these groups. The random assignment ensures that all potential confounding variables and potential outcomes have equal distribution in the groups, hence that the difference in the observed effect is causally (or randomly) associated with the applied treatment or intervention. With increasing study size, the influence from random error will decrease. However, there are growing interests in finding/identifying proper alternatives to this method because of its many limitations, such as cost, ethical aspects, and feasibility. Using observational data has been tempting for many researchers, as the data is more available, less costly and can include large observations to increase the precision of the estimates. For descriptive purposes, such observational data has been invaluable. However, when trying to abstract causal associations, several limitations have been detected through the years. (127) (128) Nevertheless, with good knowledge of how the data reflects reality, there is potential to use study designs mimicking the design concepts from randomised controlled trials to estimate causal associations also in observational studies. (68)

An often-cited quote is this: *"In God we trust, all others must have data".* (Unknown origin). This mindset has been crucial for the progression of evidence-based medicine, and health services research. Albeit, this annotation has also been added in later years: "*...and know how to use them*". In the following, I will provide the rationale for our choice of methods to use observational data to answer our causal research questions.

4.4 Study design Paper II

Our main aim in this study was to identify the effects of physician characteristics on hospital admissions and patient outcomes after assessment in out-of-hours services. In contrast to the regular GP service where patients are affiliated with a specific GP, resulting in systematic differences in the patient populations between the GPs, the out-

of-hours services could serve as an ideal setting since patients have limited choice of physician here. Nevertheless, practice variation between out-of-hours physicians could occur due to local unobserved variation in patient needs or organisational factors, such as distance to hospital or the availability of municipal acute wards. Out-of-hours services may have different patients in a small rural municipality than one in a large city. Further, patients visiting the out-of-hours setting a Friday night may differ from those visiting on a Monday night. Moreover, a Friday night patient in Oslo may differ from a Friday night patient in Berlevåg. Hence, local adaptions between primary care and specialised services may give profound differences in patient characteristics between providers. Based on this knowledge, we tried to find situations in which patient groups were comparable to each other, by systematically approaching these potential confounding factors: Patient factors such as age, sex, education level, immigrant status and previous health services use are potential confounders, as they are associated with morbidity (affecting the admission decision) and can affect both the use of health services and mortality of the patients. The mortality of the patients varies with age, geography (urban vs rural), time of the year, weekday, and the hours of presentation at the out-of-hours services. (129) Further, the characteristics of the GPs staffing the out-of-hours services likely also vary according to geography, month, weekday, and time of the day. Thus, there would most likely be some types of physicians more often assessing some types of patients, resulting in a potentially confounded association between physician characteristics and patient outcome. An example could be more experienced physicians working at night when only the patients with the most severe conditions contact the out-of-hours services. An easy way to present the relationships between these factors, is by using a casual diagram, like a directed acyclic graph (DAG). (130) In Figure 9 we show our assumptions on the relations between GP characteristics (exposure), hospital admission and patient risk of death (outcome) and patient characteristics like morbidity (confounders).

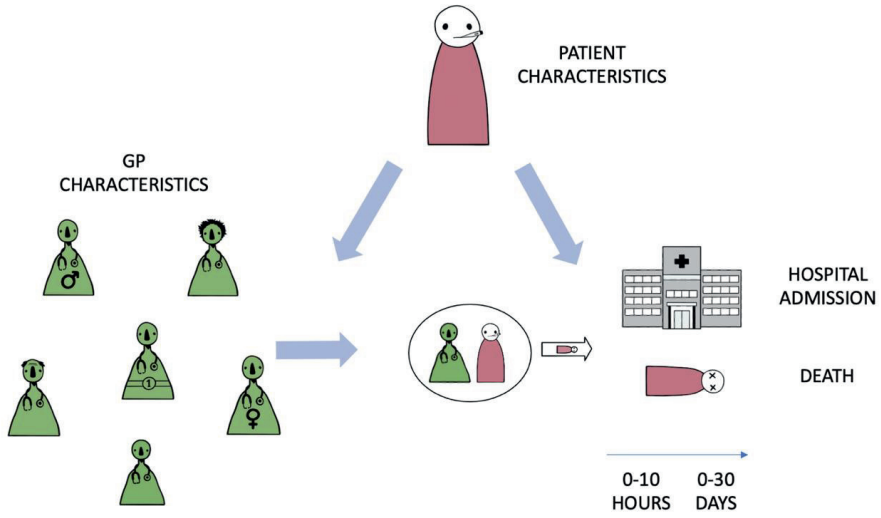


Figure 9. Directed acyclic graph showing that patient factors may be a common cause (confounder) of the characteristics of the physician they meet (exposure), unplanned hospital admission and death (outcomes).

We believed it unlikely to obtain sufficient statistically adjustments for these variations, since we may lack information on all possible confounding factors and local adaptions. (60) Therefore, we approached the question from another angle. The basic idea was to make comparisons between patients where we could arguably assume that there was close to random allocation of physicians to patients. Thus, that patients we compared, shared the same set of confounders. We took advantage of the large number of consultations and combined information in groups of patients in the same 10-years age group (based on age at the time of contact), visiting the same out-of-hours service, during the same month and year, on the same weekdays (coding public holidays as Sundays) and time of the day (in similar 8-hour partitions of the day, 16.00-23.59, 00.00-07.59 and 08.00-15.59 on weekends). By matching patients in groups based on these variables and analysing only within-group variability, we could assumably isolate/estimate the effect of the treating physician's decision and further study the associations between this effect and the different physician characteristics. (Figure 10) The close matching of patients was particularly important, since we wanted to estimate

effects from all patients in all age groups visiting out-of-hours services over several years.

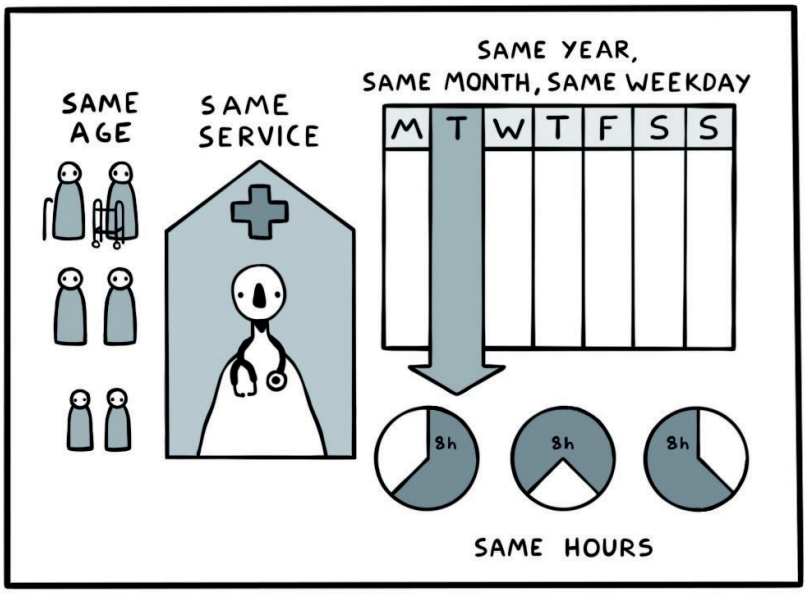


Figure 10. The matching in our study design, Paper II.

In an approach to assessing whether the admission decisions were appropriate or not, we wished to study not only the differences in admissions, but also the further consequences for the patients, associated to the GP characteristics. Hence, if a decision not to admit the patient, simply led to a postponement of admission for this patient, or if the admission was prevented. An example would be if meeting a physician with low admission proportion was also associated with a higher chance of unplanned hospital admission the following period. Further, we also wished to assess the costs from this hospital use.

Moreover, we wished to assess whether more (or less) restrictive admission practice was associated with the risk of dying. Therefore, we also included analyses of accumulated unplanned hospital admissions and risk of death in the outcome measures.

As a further measure of the appropriateness of the admissions, we included admissions for "critical conditions". (For details, please see section 4.8 Study variables). The idea here was that if the GPs' admission practice were associated with admitting patients with severe or critical conditions, e.g., high admission proportion was associated with low chance of admitting patients with critical conditions, this could point towards a higher share of inappropriate or unnecessary admissions.

4.5 Study design Paper III

For the study presented in Paper III, our main aim was to estimate the effects of being referred to the hospital from the out-of-hours services for patients where the indications for such referral were unclear. Further, we chose to concentrate on the effects for elderly patients, as these are likely more challenging to assess and highly affected by the referral decision, as they often are frailer. To increase the contrast to the regular GP setting, we chose to only include patients who are presumably unknown to the physician and staff (for details, please see section 4.7 Study population Paper III). Knowledge about the effects of an acute referral to hospital for these elderly patients with unclear need of referral, will be valuable for estimating the consequences of changing the physicians' referral threshold e.g., due to limited patient capacity in hospital emergency departments. We intended to investigate the effects on both the patients and the dynamics of the health services. Thus, our outcomes of interest were defined as health service use and mortality following the referral. For this paper, we wished to assess the more long-term effects and defined the follow-up period to be 180 days.

Again, from both previous literature and clinical experience, we know that these situations of doubt about the necessity and appropriateness of a referral are common in everyday practice. (76, 78) However, as the referral process intends to identify and select the sickest patients for a referral, obviously, the chance of being referred to the hospital after an assessment by an out-of-hours physician is associated with the chance of further health service use and death. This phenomenon is known as *confounding by indication* and will give biased results. (131) In our case, even if adjusting for all

measured confounding variables, conventional regression analyses will compare the patients selected as “healthy” with the patients selected as “sick”. Thus, they will most likely show a high association between referral and hospital use and risk of death for the patients. Nevertheless, this will not be sound evidence to suggest that patients should not be referred.

Confounding by indication could be solved by conducting a randomised controlled trial. In an imaginable RCT, we could include patients for whom their physicians doubted the necessity of a referral and randomized these patients to either referral or no referral, thus providing evidence about how a referral affects both health care use and mortality for the patients. In such an RCT, we would not include patients where the physician had already made the referral decision, as this would be unethical and could distort the outcome we wanted to study.

Instrumental variable analyses may provide results similar to RCTs in observational data by exploiting other variables with naturally occurring variation in the dataset. (132) Such variables with a known effect on the exposure of interest are referred to as instruments and can thus estimate a randomly assigned exposure level. However, the validity of the results rests upon some assumptions. Firstly, the instrument must be associated with the exposure (relevance assumption). Further, it cannot have common causes with the outcome (independence assumption), and the effect from the instrument on the outcome can only be through its effect on the exposure (exclusion restriction assumption).

The assumed relations between the exposure, outcome, instrumental and confounding variables are shown in the Directed Acyclic Graph (DAG) Figure 11. The numbers refer to the assumptions described above. Given that all assumptions are fulfilled, instrumental variable analyses can provide a causal estimate even in the presence of both measured and unmeasured confounding.

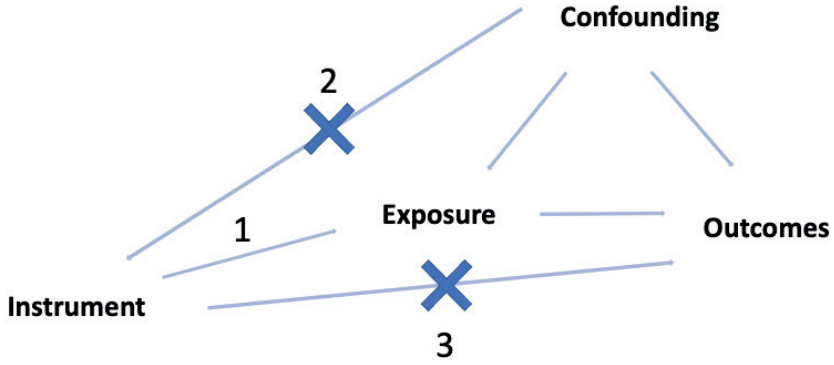


Figure 11. Directed Acyclic Graph showing the relation between the factors in instrumental variable analyses, and the three main assumptions, relevance (1), independence (2) and exclusion restriction (3).

For our study to provide valid results of causal associations, we needed to detach the reasons for getting a referral from the patient outcomes. Based on the knowledge of naturally occurring variation in physicians' referral proportion and the assumption that physicians' referral proportion and referral preference are closely related, we chose to use the physicians' preference for referral as a candidate instrumental variable. Thus, we can imagine a group of patients who are fairly similar regarding the doubted necessity of a referral. For these patients, the referral preference of the physician assessing them will most likely make the difference between getting a referral or not.

There are no perfect ways to measure the physicians' preference; however, we believed that using the proportion of out-of-hours contacts registered followed by an unplanned hospital admission within 10 hours, throughout the study period could provide a fair estimate of the physicians' referral preferences. To avoid that the referral decision made for the index patient could possibly affect the instrument, we calculated the referral preferences for male and female patients separately and applied the preference for referring male patients on the female patients, and vice versa. Based on these definitions, we assumed that confounding factors, like the index patient's morbidity, did

not affect the physicians' referral preference in the period. Further, we assumed that the referral preference did not affect the patient outcomes, other than through the direct effect on the referral decision.

The first assumption is testable by estimating the association of physicians´ referral proportion for other patients and their subsequent referrals decisions. The second and third assumptions are not directly testable, but they can be falsified, by estimating the association of the physician referral preference and their patients' morbidity and other relevant characteristics. Thus, we performed balance tests of associations between patient factors like age, education level, immigrant status, and previous health service use. The estimates of these associations were presented both unscaled, and scaled according to the strength of the instrument-exposure association. (133) Our model is presented in Figure 12:

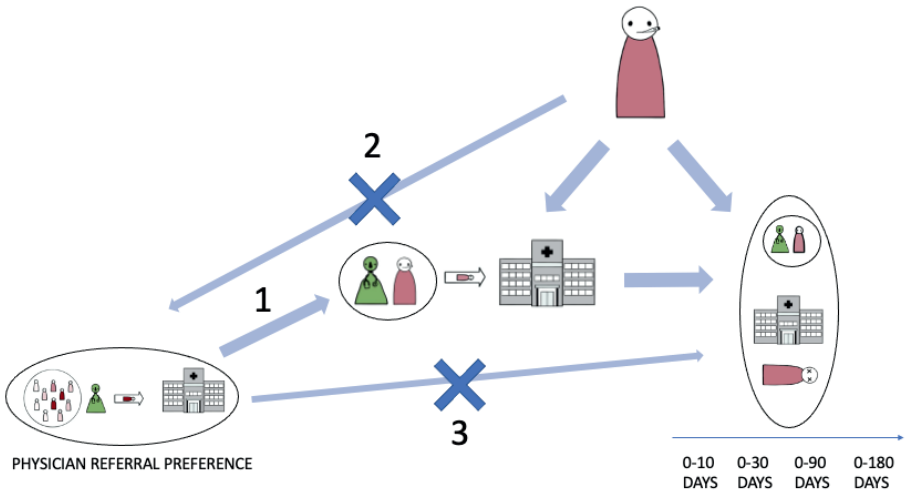


Figure 12. Directed acyclic graph showing study design in Paper III.

When applying an instrumental variable method, it is essential to be aware of the part of the population for which the effects are estimated. (134) Even though we cannot

identify the patients directly, it is valuable to have a clear view of how the model fits the data. For our dataset, we can define the four groups of patients as follows:

1) **«Always takers» -** Patients who will get a referral regardless of the physician's referral preference. E.g., patients with fractured hips or severe heart attacks.
2) **«Never takers»** - Patients who will not get a referral, regardless of the physician's referral preference. E.g., patients with minor injuries or a common cold.
3) **«Compliers»** - The patients who will respond to the instrument, that is – they will have a higher chance of being referred when seeing a physician with a higher referral preference. E.g., patients with pneumonia, vague symptoms that could imply severe illness demanding hospital treatment, but most likely is not.
4) **«Defiers»** - Patients who will get the opposite effect of the instruments. E.g., a patient who is not referred by the physician with high referral preference, or a patient who is referred by a physician with a low referral preference.

The "always takers" and "never takers" described above are patients who will never be affected by the physicians' referral preference. Based on the condition they present with, they will receive their referral (or not). The "compliers", however, are the patients who will be affected by the referral preference of the physician they meet. Hence, if they contact the out-of-hours services on a day where a physician with high referral preference is on call, their chance of being referred is higher than if they make contact on a day where a physician with a low referral preference is on call. The effect estimates from our instrument variable analyses, referred to as the local average treatment effect (LATE), will only be valid for this group of patients. (135, 136) The relevance of such effect estimates from instrumental variable analyses has been questioned, because this group of "compliers" is not always easy to identify. However, in our data, we argue that although they are not perfectly identifiable, such "compliers" will mostly be patients where the physicians have doubts about the necessity of the referral. Further, these patients are also the ones we believe will be affected by a change in the referral thresholds, e.g., if physicians are urged to reduce their referrals due to limited hospital capacity. This supports the use of physician referral preference as an instrument in our analyses.

Based on the comparison with the design of a randomised controlled trial, it has also been argued that there is a fourth assumption for getting valid results from instrumental variable analyses, namely that there are no "defiers" (monotonicity). *(134)* As described above, "defiers" are patients getting the opposite treatment of what they are randomised for. The presence of defiers in either RCTs or instrumental analyses would disturb the results and should be avoided. In our case, we assume the presence of defiers as unlikely, however not testable.

Similar to the study design in Paper II, was the possibility of confounding from a selection of patients to physicians, e.g., if more experienced physicians tended to work in larger cities, with younger patients, or during night shifts, when the patients tend to be sicker. However, for the study in Paper III, we had already selected the elderly patients, who are presumingly unfamiliar to the out-of-hours physicians and staff. Also, the only physician characteristic we needed to be randomly assigned, was the physician referral preference. Thus, we could allow for less strict matching of patients than in Paper II. However, we still believed that some matching was necessary to make the patient groups comparable: Due to the local adaptations in the different out-of-hours services, we only compared patients visiting the same services. Further, we believed the out-of-hours shifts (late shift: 16:00-23:59, night shift 00:00-07:59, of day shifts during weekends or holidays: 08:00-15:59) to differ regarding both patients' characteristics, and organisational factors, like availability of other services. Lastly, we addressed the possible time trends over the years. Additionally, since our instrument was based on the patients' sex, we compared only female patients with other female patients, and male patients with other male patients.

Thus, the matching was defined by combining information on patients of the same sex visiting the same out-of-hours station in the same 8-hour time unit during the day within the same year. For example, we compared all female patients visiting the same out-of-hours station all afternoons (16:00-23:59) in 2015. By analysing only within-group variability, we effectively controlled for all confounding that was constant within each group. To avoid the effect of possible patient selection in situations where two or more

GPs were on-call at the same time, we used the weighted average of GP characteristics within each 8-hour time unit in each service. (Details of the study design are presented in Supplementary Paper III).

4.6 Study population Paper II

In Paper II, the study population comprised all contacts with the out-of-hours services in Norway from 2008-2016, registered in the Control and Payment of Health Reimbursement Register. However, to make sure we included only contacts relevant for answering the research question, we made some exclusions: Since the linkage between the physicians and the physician characteristics required the unique physician identification numbers we only included contacts with identifiable physicians claiming reimbursement. There could be various reasons for omitting physician identification number, like erroneously filed claims; however, we believe claims without this number primarily to come from physicians with fixed salaries from the municipality as such claims are often sent from the municipality responsible for the out-of-hours service. In Norway, fixed salaries are mostly provided for night shifts, however some places also for late shifts or weekends. (137)

Further, we limited the patient population to include only contacts performed during out-of-hours shifts, i.e. between 16:00 and 08:00 on weekdays, and from 08:00-16:00 on Saturday, Sunday and public holidays. We included all claims containing the specific codes for out-of-hours work where physicians assessed patients, face-to-face or by telephone (codes: 2ak, 2fk, 1ak, 1bk, 1g). We excluded claims containing codes for home visits made by physicians (codes: 2nk, 11ak, 11nk, 21k). To obtain complementary information about the treating physicians´ characteristics such as age, sex, and specialist status, we included claims only from physicians working as regular General Practitioners during office hours, who were registered in the Norwegian general practitioner register. Further, we excluded claims from contacts where the patients' regular GPs were present at the out-of-hours services since such contacts could have been arranged between the patient and the GP.

We believed the effects of the GP characteristics to differ in different patient age groups. We hypothesised that the youngest and oldest patient groups more often use the out-of-hours services for urgent health conditions where the clinical choice of immediate hospital admission is likely to be particularly important. Therefore, we chose to perform separate analyses for three age groups of patients; 0-10 years, 11-69 years, and above 70 years of age, emphasising the youngest and oldest groups.

4.7 Study population Paper III

For Paper III, we used the same data material as for Paper II, but we chose to study the older part of the study population, defined by the OECD as patients 65 years and older. (138). To ensure we included only contacts from out-of-hours work, we made the same inclusion criteria based on reimbursement codes and time of consultation. However, we included all contacts where the physician identification number were provided, and thus not only those registered in the General Practitioner register. Further, to increase the contrast to the context of the regular GP, we wished to study patients who were formerly unfamiliar to the out-of-hours physicians and the staff. Hence, we excluded contacts where one or more physicians on call had missing id, where patients had been assessed by the same physician previously during the data period from 2006-2016, where the patient met his or her regular GP, and from patients without a registered regular GP. As frequent attenders to out-of-hours services will most likely be known to the staff, (139) we excluded all contacts from patients with > 4 consultations per year (90-percentile). Patients who were registered with date of death before the date of contacts, were also excluded. The flowchart presented in Figure 13 shows the number of contacts and patients in the separate inclusion steps.

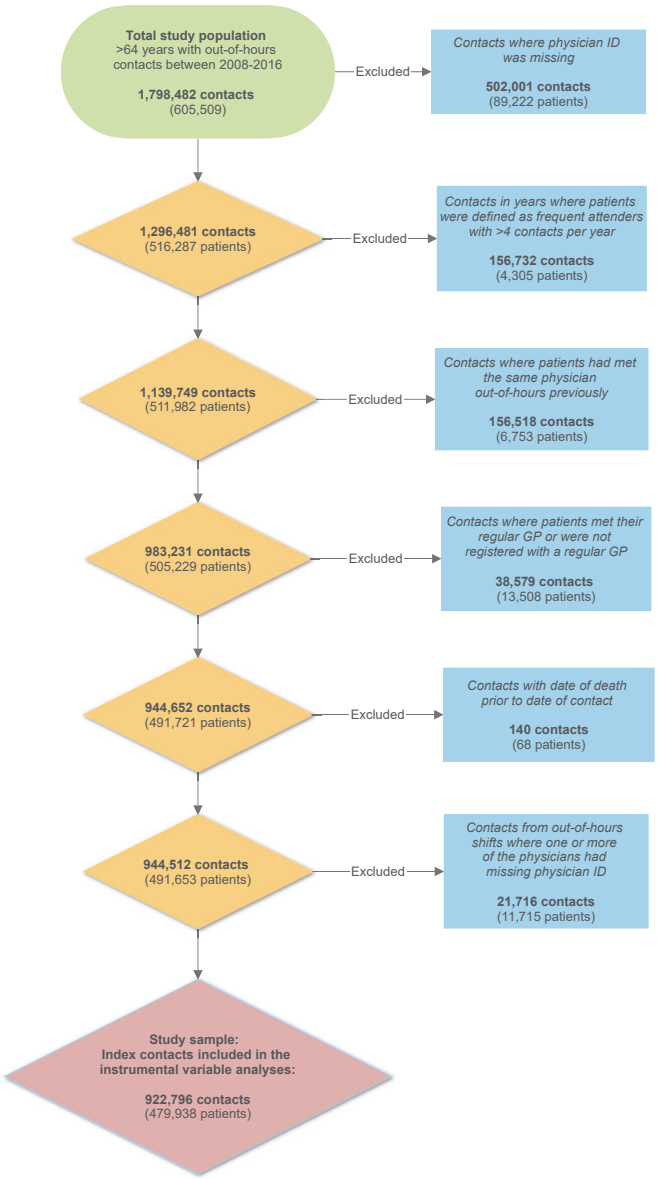


Figure 13. Flowchart, Paper III

4.8 Study variables Paper II and III

For both papers, we used the setting of patients assessed by physicians in the out-of-hours services, as defined above. In both Papers II and III, each contact with the out-of-hours services was initially defined as an index contact; however, for Paper II, we only included the first out-of-hours contact if the same patient was registered with more than one contact during the same time unit. This choice was made based upon defining the exposure as the characteristics of the first GP they met during the out-of-hours shift. For Paper III, index contacts were defined as described above.

GP characteristics

In Paper II, the GP characteristics were used as exposure variables. We defined the five exposure variables as the GP age, sex, specialist status, out-of-hours experience the previous two years, and the out-of-hours admission proportion the previous four months. (**Figure 14**)

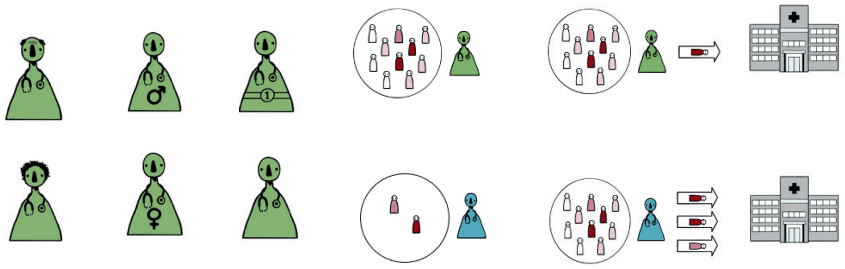


Figure 14. GP characteristics: Age, sex, specialist status, out-of-hours experience the previous two years, and admission proportion the previous four months.

GP age, sex and specialist status were based on information provided in the regular general practitioner register. Age at the time of each contact was based on the GP's year of birth, and specialist status was defined as specialist for all contacts following the eventual date of approval of specialisation in general practice. We defined the variable *out-of-hours experience the previous two years* based on each GP's number of out-of-hours contacts the previous two years. This variable was dichotomised and defined as

low if the number of contacts were less than 200. We defined the variable *GP admission proportion* based on the proportion of out-of-hours contacts followed by an unplanned hospital admission within 10 hours of the contact. We further divided the group of GPs into four equal groups (quarters) based on the value of this variable, allowing for comparison between the highest and the lowest quarter.

Immediate unplanned hospital admission / referral to hospital

In Paper II, we chose the primary outcome to be an unplanned hospital admission, following an assessment by an out-of-hours physician. Unfortunately, we had no available data on the actual referral process, as referrals are not registered in the patient registers. However, we believed it highly likely that patients with an acute referral to the hospital from the out-of-hours service would be registered with an unplanned contact or admission in an emergency department within ten hours after the out-of-hours contact. We chose the ten hours time span, as this would allow the patient some time for transport between the out-of-hours service and the hospital; however, not long enough to include a new assessment the next day. This is not a perfect measurement for the outcome of interest but based on our own clinical experience from working in the out-of-hours services, we believed it would be close enough to reflect the reality for most emergency contacts or admissions to hospitals in Norway. We also limited the hospital admissions to those made with a code for acuity, to exclude planned hospital admission coincidently occurring on the day of the out-of-hours contact. We recognize that being registered with an emergency visit to the hospital is not always equal to being admitted and that this may be a potential limitation from this definition.

The same definition of this variable was labelled “referral to hospital” and used as the outcome variable in Paper III.

Unplanned admission for critical conditions

In the quest of trying to investigate the consequences of the variation in unplanned hospital admissions for Paper II, we chose to define a modified version of the primary outcome variable to include only hospital admissions for severe or critical conditions.

We hyptothesised that if the excess admissions made by the physicians with the highest admission proportions were mainly admissions for less severe conditions, this could suggest that these admissions were less necessary or appropriate. The tentative diagnoses made by the out-of-hours physicians for the reimbursement claim would have been valuable for this purpose; however, many of these ICPC-2 diagnoses are unspecified and thus not valid for this purpose. (40) Therefore we defined "admission for critical conditions" as an unplanned hospital contact resulting in at least one discharge diagnosis from ICD-10, indicating a severe or critical condition, with a clinical consensus of the need for hospitalisation. The selected ICD-10 codes are presented in Supplementary material for Paper II and also listed below:

List of discharge ICD-10 diagnoses used to define "critical conditions"

S06 Intracranial injury (excluding S06.0 Commotio cerebri)
S72 Fracture of femur
I21 Acute myocardial infarction, unspecified
I22 Subsequent myocardial infarction
I26 Pulmonary embolism with mention of acute cor pulmonale
I63 Cerebral infarction
A41 Other Sepsis
K35 Acute appendicitis
K80 Cholelithiasis with cholecystitis (excluding S80.2, S80.5 and S80.8 which do not include cholecystitis.)
K56 Paralytic ileus and intestinal obstruction without hernia

The following criteria were defined based on discussions in the project group and on clinical experience and were used to define the original list of ICD-10 diagnoses for "critical conditions":

1) There is a professional consensus that these conditions, as the main rule, need treatment in hospitals (in the Norwegian health care model).
2) If a patient presents with one of these conditions in an out-of-hours service, it will most likely have negative consequences not being admitted to the hospital.

3) The conditions must not necessarily be easy to recognise when presented in the out-of-hours services, but it should lead to hospitalisation if they are recognised.

The selected diagnoses were presented and discussed in a forum with both practicing and academic researchers and adjusted accordingly.

Hospital admissions the following 30 days

In Paper II, other outcome variables were hospital admissions in the 30 days after the index contact. This outcome was defined as two variables: 1) Cumulative incidence of unplanned hospital admissions and 2) Cumulative costs from unplanned hospital stays starting within 30 days after the index contact. The first variable was defined as the number of unplanned hospital visits started within the first 30 days after each index consultation and was meant to capture if the initial decision of not admitting the patient only led to a short postponement of the admission, e.g., that the patient was admitted by either the same or another physician after all. We believed this outcome to be important and "new" in such research questions. The second variable was included both to show the economic consequences of the variation in admissions and to reflect the degree of hospital treatment. To some extent, it would also reflect the severity of the patient condition. We based this variable on the diagnose related group (DRG) points generated from each hospital visit. DRG points are made as part of the hospital financial system and are based on the patient's diagnoses, combined with the length of the hospital stay, the procedures and the level of care provided. (140) Outpatient contact will generally generate lower DRG points than inpatient treatment. Similarly, a stay in the intensive care unit will generate higher DRG points. We used these DRG-points to calculate the costs for each hospital visit started in the 30 day following each index contact. By combining the DRG value based on the 2016 unit price, (140) with the average Euro exchange rate from 2008-2016, we defined one DRG-point to value at € 5075. The cumulative incidence of hospital stays, and the hospital costs were presented per 1000 out-of-hours contacts, to show the consequences in a greater context.

Health services use

In Paper III, our primary outcome variables were health services use following the index contact with the out-of-hours services. We defined this as the number of days with registered contacts in three levels; 1) Primary care physicians, 2) Hospital stays, and 3) Outpatient clinic visits. Health service use in these three levels were measured for the four time periods following the index contact: 1) 0-10 days, 2) 0-30 days, 3) 0-90 days, and 4) 0-180 days. This definition gave us 12 outcome variables of health service use for each index contact and allowed us to carefully investigate the dynamics between primary care and specialised services in each patient trajectory. The variable "contacts with primary care physicians" included both visits to general practitioners working daytime and out-of-hours.

Mortality / Risk of death

In both Papers II and III, we assessed patient safety as the mortality rates for defined time periods following the index contact. For Paper II, we defined the 30-days risk of death by information on death registered in the Cause of Death Registry within 30 days of each index contact. In Paper III, we wished to assess the hazard ratios for death within the different time spans from the referral decision. As outcome variables, we therefore included deaths registered in the Cause of Death Registry within 10, 30, 90 and 180 days following the out-of-hours contact.

Other variables

Modal municipality

There is extensive collaboration for organizing the out-of-hours services between municipalities leads to shifting geographical locations of many out-of-hours services, especially in the scarcely populated municipalities. The same physicians may often serve the population from several municipalities within the same period. To avoid possible confounding from comparing out-of-hours activity in, for example, a large city with a rural municipality, the matching of patients in the study designs in Papers II and III were based on patient contacts from the same out-of-hours service. In Norway, all municipalities have a unique municipality code. Unfortunately, the Control and

Payment of Reimbursement Register claims lack information about the actual geographical location where the out-of-hours contact took place. However, we had information about the patients' municipality of residence from Statistics Norway. We assumed that in most cases, patients visit the out-of-hours services in their home area. Hence, to define the location of the out-of-hours services, we used the municipality codes´ modal value among patients within physicians per week. As a result, the municipality where most of the patients (within one week, seeing the same physician) lived was defined as the municipality where the contacts for this physician took place that week. Thus, this could change through the time periods in the data material. If the value was the same for two or more municipalities, we chose the lower code, which often represents a larger municipality.

Education

Patient education level at the time of each contact was calculated based on education information from Statistics Norway. For use in Papers II and III, we dichotomised this variable, and defined it as "low" for patients who had completed less than 13 years of school.

Immigrant status

Patient immigrant status was also based on information from Statistics Norway and dichotomised. Patients were defined as immigrants if they were registered as immigrants or Norwegian-born to immigrant parents.

Previous health care use

As measures of the patients' morbidity, we chose to make variables for the patients' health service use the month preceding each index contact with the out-of-hours services. We chose to separate the use into the following four variables: Unplanned admission, elective admission, outpatient specialist clinic visit and primary care physician contact, the three first based on data from the Norwegian Patient Register, the latter on data from the Control and Payment of Reimbursement Register. These variables were dicothomised and given the value 1 if the patient was registered with the health service use in the previous 30 days before the index contact.

Charlson Morbidity Index

Another, and more widely used measure for patients' morbidity, is the Charlson Morbidity Index. (141) We made a variable for this index for all patients, based on the diagnoses registered in the Norwegian Patient Register the month before each index contact with the out-of-hours services.

4.9 Software

For Paper I, we used Ecxel to aid in the Systematic Text Condensation. For Papers II and III, all statistical analyses were performed using Stata version 15.1.

4.10 Statistical analyses Paper II

For the statistical analyses in Paper II, we used linear and logistic regression with within-matched-group estimators (xtreg with FE option and clogit in STATA) to estimate the associations between GP characteristics and the patient outcomes. We matched all patient contacts in groups (described under Study design) to address possible unmeasured confounding. Based on the assumptions presented above, we assumed that all patients we compared in each group, shared the same set of confounding variables regarding location, organizational factors, and systematic variations in patient factors. All multivariable analyses were performed using the grouping variable to estimate within-group variability. In such analyses, only groups with differential exposure and outcome are included, as groups with non-differential exposure or outcome will not contribute to the effect estimate. For example, in the analyses of GPs' sex and chance of hospital admission, groups comprising only patients assessed by male physicians (exposure) or where no patients were admitted within the follow up period (outcome) were not included. Hence for each of the different exposures and outcomes, different groups were included, contributing to a varying number of observations in each of the analyses. The out-of-hours experience variable was dichotomised. For the physician history of unplanned admissions, we compared the highest and the lowest quarter. Immediate unplanned hospital admissions and 30-day risk of death were analysed using a within-matched-group estimator with conditional

logistic regression (clogit command in Stata), while the 30-day unplanned hospital admissions and costs (for hospital stays starting 0-30 days after an index contact) were estimated using a within-matched-group estimator with linear regression (xtreg, fe in Stata). All estimates were adjusted for patient sex, age and age squared.

4.10.1 Assumptions and sensitivity analyses Paper II

An important assumption for the validity of our results was the independence between possible confounding variables on the patient levels and the GP characteristics. This assumption cannot be proved but can be falsified if associations between patient characteristics and GP characteristics are found. Hence, we performed balance tests associating the GP characteristics with patient pre-treatment characteristics.

4.11 Statistical analyses Paper III

In Paper III we used instrumental variable analyses to estimate the causal associations between referral from out-of-hours services, and the patients' further health service use and mortality. As for Paper II, we assumed that all patients we compared in each group, shared the same set of confounding variables regarding location, organizational factors, and systematic variations in patient factors. We used the within-matched-group estimator for instrumental variable regression (ivreghdfe in STATA) to study the effects on further use of health services and ran separate analyses for all exposures and outcomes. (142) Also here, only groups with differential exposure and outcome were included in the analyses. We adjusted for month, weekday, patient age, age squared, and follow-up time (number of days at risk), in addition to the within group estimation. Mortality was analysed with a within matched group estimator using stratified Cox analyses, and a two-sample IV estimator (the delta method). (143) As a comparison, we used multivariable adjusted linear regression analyses to assess the same associations. Here, we adjusted for year, month, weekday, consultation hour, patient sex, age, age squared and follow-up time (number of days at risk) for each index contact. We used Cox regression with time from index consultation as the time axis to investigate the association between referral and mortality. Precision was evaluated with 95%

confidence intervals (CI) with robust standard errors clustering on physician and patient ID. (142)

4.11.1 Assumptions and sensitivity analyses Paper III

As described in section 4.5 Study design Paper III, valid results from instrumental variable analyses rest upon some assumptions The relevance assumption was tested as the instrument-exposure association, where the F-test should at least give values > 10. (144) However, as any violations of the IV assumptions will be amplified with a weak IV-exposure association, a substantially higher F-value would be preferable. (145) The independence assumption can be supported by analysing the association between the instruments and variables known as confounders to the exposure-outcome associations. Therefore, we analysed the associations between physician referral preference and patient characteristics such as age, immigration status (yes/no), education (completed less than ten years or more than 13 years), and health service contacts 30 days before the index contact, the latter as primary care physician contacts, planned and unplanned hospital admissions, and visits to outpatient specialist clinics. Additionally, we used discharge diagnoses the from previous hospital stays divided in main chapters (ICD10, Chapter IX Diseases of the circulatory system, and Chapter II Neoplasms). These associations between potential confounders and the instrument were also presented as scaled based on the strength of the instrument-exposure association.

4.12 Ethics

The study in Paper I did not obtain any patient information and was approved by the Regional Committees for Medical and Health Research Ethics (REK) (2016/2158/REK-midt) and the Norwegian Data Inspectorate. All participants signed a written consent to participate. They were informed that participation was voluntary and that they had the opportunity to withdraw from the study at any time. The study information provided in the invitation upon consent to participate described how data would be stored securely and de-identified as soon as the process allowed for this. (See section 4.13 Anonymity).

The studies in Paper II and III were a part of the larger project “Health care services under pressure - Consequences for patient flows, efficiency and patient safety,”

approved by the Regional Committees for Medical and Health Research Ethics (2016/2159/REK midt). Participant consent was not required in these studies.

Our research project group is affiliated with several patient organisations, and the projects and results have been presented and discussed with representatives from these organisations during the project period. We have also sought advice on the appropriate dissemination of the results. We believe that the scientific contribution from this thesis can provide valuable insights to the health care system and be valuable for policymakers in the planning and organising of a better use of the health services resources.

4.13 Anonymity

In studies involving information from humans, the identity of the participants should ideally be anonymised so that they will never be traceable from the data. In qualitative interview studies, this is known to be hard to achieve. (146) In the study presented in Paper I, we approached this with the following measures, in line with the approval from the Norwegian Data Inspectorate: The raw data (audio recording) was stored on a secure/encrypted server until the transcripts were secured and proofread. During the transcription process, all names, dates, places, and people mentioned were coded with numbers and capital letters, and an encryption key was made accordingly. The audio recordings were then deleted. Participant consent and information schemes were stored securely, separated from the data material. We used the encryption key to link the participants to data material/quotes, only accessible for me as the first author. This encryption key is now deleted.

The GP workforce in Mid-Norway is relatively limited, and people with knowledge of this group of GPs could potentially recognise participants if too many details were provided, such as the descriptive characteristics. We paid particular attention to this when we presented descriptives and selected the illustrative quotes. The translation of quotes from Norwegian to English further contributed to the anonymisation.

In the studies presented in Papers II and III, we used the register data link, approved by REK as described above. The data storage and handling were in line with the approvals from register owners REK. Details are provided in the project Data Protection Impact Assessment (DPIA). For these studies, we did not know the identity of participants (neither patients nor physicians), as they were all given a serial number upon provision of the data. Hence, each patient contact was deidentified and provided a serial number to link the information from the different registers. However, this implies that information such as patients' year of birth, time of contact, municipality and diagnoses were linked, with a potential for misuse if not secured. Therefore, these data were stored securely on an encrypted server, with a two-step identification process to access. We have also ensured that all results were presented on superior levels, not to be identifiable.

5 Main results

In the following section I will present summaries of the main results in the three papers. For closer details, please consult the result sections in the respective papers. Further, I will present a selection of supplementary analyses from Paper II and III.

In Paper I, our analysis identified three main themes: (1) Heavy and increasing workload – more trend than fluctuation?; (2) Explanations for high workload; (3) Consequences of high workload. Our findings showed that both GPs and their co-workers experienced heavy and increasing workload. The suggested explanations varied considerably among the GPs, but the most commonly cited reasons were legislative changes, increased bureaucracy related to documentation and management of a practice, and changes in patients' expectations and help-seeking behavior. Potential consequences were also perceived as varying, especially regarding consequences regarding GPs' health and motivation, as well as the recruitment of new GPs. However, the GPs were aware of their gatekeeper function and made efforts not to increase referral of patients as they recognised the potentially added burden of workload on the health services.

In Paper II we found substantial variations in GPs admission practices. Patients under the care of older and male GPs had fewer immediate unplanned hospital admissions, and these effects were most evident for elderly patients and children. However, the effects on cumulative 30-day unplanned hospital admissions and costs were small. The GPs' prior admission proportion was strongly associated with both immediate and 30-day unplanned hospital admissions, also this particularly for elderly patients. Notably, higher prior admission proportion was also associated with admitting more patients with critical conditions. There was little evidence of any associations between GP characteristics and 30-day risk of death, however for patients assessed by GPs with low out-of-hours experience there was a slightly increased 30-day risk of death.

In Paper III we found that for elderly patients, whose referrals were attributable to their physicians' threshold for referral mean length of stay in hospital was 3.25 days (95% CI 3.08-3.42) within the first ten days, compared with non-referred patients. Such referrals

also increased six months use of outpatient specialist clinics and primary care physicians. Importantly, patients with referrals attributable to their physicians' threshold had a substantially reduced risk of death the first ten days (HR 0.53, 95% CI 0.29-0.92), an effect sustaining through the six months follow up period (HR 0.70, 95% CI 0.50-0.95).

5.1 Sensitivity analyses Paper II

Balance test of confounders

Our study design depends on strong assumptions of independence between the GP characteristics and the patients' characteristics prior to the index contacts. These assumptions cannot be validated, however, they can be falsified, if regression analyses show associations between GP characteristics and patient factors known to confound the exposure-outcome association. To explore these possible associations, we performed regression analyses with a set of potential confounders. We found weak or no associations between the GP characteristics and the possible confounders, supporting our assumptions of independence in our study design. Results are presented in Supplementary table 1 (Paper II).

Multiadjusted regression analyses for primary outcomes

In the main regression analyses, we adjusted only for patient factors. However, we made sensitivity analyses for all primary outcomes, where we adjusted all exposure variables for each other. This did not substantially affect the results, and in particular, the strong effects of the GP prior admission proportion were largely unchanged. The results are presented in Supplementary Figure 1 (Paper II).

GPs vs non-GPs

For our main analyses, we chose to include only patient contacts with active GPs working in the out-of-hours setting, as the pressure in the health services may threaten the contribution from GPs in the out-of-hours services. The GPs contribute with about half of the contacts in the out-of-hours services. The rest of the contacts are provided by non-GPs staffing the out-of-hours services. These comprise a heterogenous group of

physicians (e.g. hospital physicians in various specialities, interns, PhD fellows and locums). Even if studying differences between the performance of GPs and non-GPs was not in the scope of this study, we performed sensitivity analyses to ensure to detect any striking differences in the referral practices between the groups. The analyses showed a slightly higher OR for immediate unplanned admission after contact with a non-GP compared to a GP. Results are presented in Supplementary Table 3 (Paper II).

Ten-year age groups

In the main analyses, we chose to present the results for three main age groups, with main emphasis on the youngest and oldest patients. However, as sensitivity analyses, we showed the results also for all ten-year age groups. We found that the results are consistent over all age groups, however the estimates are slightly higher for the oldest patients. Results are presented in Supplementary Table 4 (Paper II).

5.2 Sensitivity analyses Paper III

As for Paper II, the study design presented in Paper III depends upon strong assumptions of independence between the physician referral preference and the systematic differences. Further the validity of the results depends on a valid instrument. To support our assumptions if independence and the validity of the instrument, we performed some sensitivity analyses presented below.

Balance tests/confounder analyses

For the study presented in Paper III we performed balance test of confounders to support our assumption of independence between physician referral preference as the instrument and the patient characteristics for the index contacts within our defined groups. We made variables for the health services use prior to the out-of-hours contact for all patients. We also made variables based on being discharged with specific diagnoses; 1) ICD-10 Chapter IX, Diseases of the circulatory system and 2) ICD-10 Chapter III, Neoplasms (Malignant diseases). Together with variables for patient sex, age, education level and immigration status, these variables were used in analyses to test whether patient characteristics were associated with the referral preference of the out-

of-hours physician assessing them. The results presented in Paper III, Supplementary Table 1 are shown both unscaled and scaled according to the strength of the instrument-exposure associations. The overall impression of the results is that the confounders, i.e., the systematic variation in potential confounders are well balanced in our design, supporting our independence assumption.

However, for this study we chose to show the distribution of confounders also for the multivariable adjusted regression analyses. The results are presented in the same table, and show a higher likelihood of previous healthcare use, and higher age for patients assessed by physicians with higher referral preference. This unequal distribution of confounders supports our use of an instrumental variable approach, as it indicated a confounded exposure-outcome association with conventional regression analyses even after adjusting for measured confounders.

6 Discussion

6.1 Main findings

This thesis presents three studies examining different aspects of the complex dynamics in the health services, emphasising the effects of the increasing demands on health services and how measures to reduce pressures in one part of the service may reflect as increasing pressure in other parts or even compromised patient safety. Our results provide insight into how the general practitioners perceive and tackle workload, their reflections about how this can affect their gatekeeper role, and the sustainability of the regular general practitioner service, currently an essential part of the gatekeeper system. The general practitioners perceive workload as high and increasing, listing a striking variety of reasons, and they express concerns about the sustainability of the general practitioner scheme. Through novel methodological approaches to the use of register data, we studied the primary care physicians' gatekeeping role, emphasising the impact from their characteristics and thresholds on referring patients to the hospital. Our findings reveal substantial effects from GP characteristics, their threshold for referral and admission of patients seemingly of most importance, with consequences for the demands on the health services and directly for patients, both in the means of unequal distribution of services and patient safety.

6.2 Methodological considerations – strengths and limitations

The studies presented in this thesis comprises different methodological approaches to answer our research questions. With access to comprehensive health register data, we acknowledged the value of combining quantitative and qualitative methods to make the research questions as relevant as possible. Even if our researcher group comprise many professions including physicians with relevant work experience, we found it most important to explore the field directly to assess updated knowledge. In the following, I will elaborate and discuss the different methods, including their strengths and limitations. I will also reflect upon my own prerequisites as a physician, and how this may have affected our research and findings.

Both in the use of qualitative and quantititative methods, the validity of the research and findings are of cruicial matter. Validity is often divided into external validity, which points to the usefulness or relevance of the findings, and to internal validity, referring to the degree of trustworthiness, as in a lack of systematic error. External validity further relies upon the internal validity. Systematic error, also referred to as biases, can be explained by the data selection process, the measurement of study variables or by confounding. In contrast, random error is defined as the variability in the observed data due to chance or unobserved causes. Both of these should be minimised to optimise the validity of a study. How to obtain validity differs between the two methodological approaches. In the following I will discuss some important aspects of validity in relation to our chosen methods. I will start by discussing validity of our qualitative study (Paper I), followed by the quantitative studies (Paper II and III). The external validity of our findings will also be discussed in later sections regarding their relation to existing literature and implications for future research.

6.2.1 Paper I Transferability and reflexivity

External validity in quantitative research is often referred to as generalisability, namely if the findings can be generalised to other populations than the source population. In qualitative research, the term transferability is more appropriate, as the aims are more directed at describing nuances rather than similarities. (120) Transferability can be defined as to which degree the findings can provide new insights that others can benefit from, also in other contexts than in the study setting. Transferability highly relies upon the reflexivity of the researcher. (120) This can be regarded as the researchers´ reflections upon their preconceptions and their role in the planning and conduction of the study, the collection and analyses of the data, and the interpretation of the findings. In our qualitative study, my preconceptions likely have influenced my motivation for conducting the study, my attitudes towards the participants, and the interpretation of our findings. Through working as both a general practitioner and a hospital physician, I have gathered various experiences of workload. In some means, periods of high workload has given me a sense of being more efficient and powerful in working, however in other means, and especially with persistent high workload, I have felt insufficient and incompetent for performing my tasks. Further, I felt high workload

could to some degree influence my decision making, for example, regarding referrals of patients. These perceptions fed my interest in studying how other health care personnel perceived and were affected by workload, and the potential consequences this phenomenon could have for the health services, and consequently for the patients. This interest served as a base for the formulation of the research question. However, I was highly aware of these preconceptions' potential influence on the qualitative research process. To reduce such influence, I tried to actively distance my own experiences and feelings from the data material, in all parts of the process, partly through reflecting upon the similarities and differences between these two. Furthermore, my fellow researchers' preconceptions have impacted the research process, and we actively approached this by discussing our preconceptions and observations throughout the whole process in the group. Although striving to be reflexive, our own experiences and attitudes still serve as limitations for the findings.

We were four researchers with work experience from general practice, one of whom currently worked as a regular GP. We all had personal experience with periodically high workloads and expected this to be the case for our participants. These expectations may have affected the interview guide and our follow up questions during the interview settings. However, the interview guide was adjusted as we observed that the participants consequently went over from engaging in periodically high workload to a steady trend of increasing workload.

Further, during the interviews, our role as moderators may have been influenced by our professional background. We may have expressed understanding and empathy when the participants told us about their experience with workload. If we had responded with, e.g., curiosity, the further discussion could have been somewhat different. As an example of how the researchers could affect data collection, I observed and reflected upon that when the economist acted as moderator, she followed up on different perspectives than what I felt would be natural. Such observations were discussed in the group. Our interdisciplinarity, could also serve as a strength in our process, as it challenged our different preconceptions. It may have counteracted the unification of the findings by contributing to variation in the interpretation of the data.

The process of planning the interviews could probably have been more structured and based on methodological considerations. As the topic we wanted to study was closely related to not being available for participation, we had worries about excluding important informants if we were too rigid on the premises. Hence, we emphasised to appear as flexible as possible for the participants. This resulted in the different constitution of groups in all the interview sites and some deviations from the recommended group size of 5-8 participants in the focus groups. However, we continued the interviews until we agreed that we had reached saturation, and the diverse interview settings were perceived as not substantially negative in this setting.

By following a more structured approach to the inclusion of participants and staying more true to the recommended number of participants in each group, we may have provided results easier to compare with existing literature. The large diversity in the conduction and reporting of qualitative research has been problematised. (147) On the other hand, by not including the participants from the small groups, we would have possibly lost invaluable insights and perspectives on the topic. As a reflection from writing this thesis, I feel we could have found another term than focus groups for these interviews, as the wording may be misleading to the readers.

Our choice of varying the combination of professions in the focus groups had several reasons. One was to increase the number of eligible participants for the study, especially in practices with few physicians. Further, we wished to allow participants to enlighten the topics and phenomena from different perspectives. In the community within a GP practice, GPs and coworkers are close colleagues who mostly know each other well and rely on each other daily. Still, one must be aware of the skewed relations, as the GPs often serve as the coworkers' employers, both financial and organisational. In a focus group setting, such relations could lead to self-censoring or a higher level of agreement. However, a comprehensive understanding of each others working conditions can help enrich and nuance the discussion of the topic. We observed such effects in our focus groups, as the coworkers and the GPs could provide and discuss different aspects of the same phenomenon, e.g., such as how the workload and the dynamics in the patient

flows were affected when one or more of the GPs were absent from work, or during low-activity periods like in the summertime. The GPs experienced this to result in higher workload having to compensate for their absent colleagues, while the coworkers had less workload in such situations. However, the coworkers perceived higher social and emotional strains in these situations because they had to turn down patients who needed to see a GP.

Some general limitations of collecting data through focus groups and individual interviews should be considered. Through such methods, the researchers will only be provided with the thoughts and information that participants are willing to share. It is possible, particularly in focus groups, that the participants hold back on details, which can leave them with a negative appearance. In our case, discussing workload, we experienced that the participants shared openly stories and details that could be sensitive to them, for example, how workload made them make bad decisions and perform actions they were not proud of. In these situations, their colleagues acted supportively, which contributed to further sharing such stories. The fact that health care personnel daily rely on confidentiality may have contributed to this feeling of confidence in the interview setting. However, we must be aware that there still might be a high degree of self-censoring. If there had been a substantial degree of self-censoring in the focus groups, the individual interview settings could have allowed the participants to share their thoughts in more confidentiality. The stories and experiences shared in the two settings did not differ considerably, supporting that self-censoring was not substantial.

Further, the timing of the study, with the concurrent media focus on workload in general practice, may have influenced the participants' engagement in the topic, the threshold for expressing negative opinions, and their reflections and perceptions on how workload affected them and their practice. This was likely an advantage regarding the recruitment of participants for our study, and we believe it might have accentuated the opinions expressed by the participants. In hindsight, it would have been interesting to include questions on how the GPs could reduce their workload and go further into the dynamics of how the GPs perceive their own contribution to workload both in general practice and in other parts of the service.

Our study was conducted in Mid-Norway, and although including both rural and central GP practices, there may be systematic differences from other parts of Norway and health care systems in other countries, limiting the transferability of our results. However, some of the enlightened mechanisms, like the diversity of causes for the increasing pressures, may be valid for many other parts of Norway and other countries, especially with similar gatekeeper systems. As many of the reasons discussed by the participants were specific for Norway, like changes in legislation, and sick leave certification, these would most likely not apply for primary care settings in other countries. However, the experience of increasing demands from the population, increasing bureaucracy and documentation needs are similar to findings from England and may well be transferable to other health care systems. (105)

6.2.2 Paper II & III Precision, validity and study design

For the studies presented in Papers II and III, we used the comprehensive register data linkage. As described in the methods section, using observational data to answer causal research question requires close attention to possible pitfalls of systematic error, such as selection bias, information bias and confounding. The comprehensive health registers and the extensive use of the 11 digit personal identification number in Norway are considered a gold mine for researchers. The data is collected, organised and ready for use, there is no need for patient consent as it is regulated by legislation. (15) Nevertheless, there are also some potential drawbacks. Relying on register data collected for other purposes than research further requires caution, and brings some potential limitations. In the following, I will discuss some aspects important for the presicion and internal validity in our quantitative studies.

Precision

Precision can be defined as the lack of random errors. (148) However, there are several different traditions of assessing and interpreting presicion in observational studies. Our approach is based on an estimation tradition, where we avoid null hypothesis testing,

and rather emphasise to assess presicion using the confidence intervals of the estimates. (149) In this way, we use the range of the confidence intervals to assess where the likely effect estimate lies, presuming our model assumptions are valid. Thus, the assessment of presicion also relies on the validity of the model.

A main strength of our register data linkage is the large amount of data, reducing the influence of random errors on the associations under study. This is reflected as precise estimates, with narrow confidential intervals. Overall, for the main analyses in both our studies, our estimates are relatively precise. However, the study design with matching of observations to avoid confounding, comes at the price of precision, as the groups with non-differential exposures or outcomes are omitted from the analyses. Thus, some of the results from analyses with less frequents outcomes, like from the analyses of risk of death associated with physicians out-of-hours experience in Paper II are less precise, with confidence intervals rangeing from both substantial positive to negative effects. Also for our secondary outcomes, like critical conditions and 30-day risk of death (Paper II), the results were less precise. This may have been improved by a or by longer follow up time, or by less close matching; however, with the risk of introducing confounding.

In Paper II, we made a variable based on each GP's previous admission behaviour, calculated as the proportion of out-of-hours contacts leading to unplanned admission for patiens assessed by the physician during the previous four months. Using only a relatively short period of time decreased the presicion affecting the presicion of the estimates of the risk of death. For Paper III, we chose to use more information for each physician, namely by including all out-of-hours contacts through the whole study period, for each patient sex separately. This resulted in a substantially increased statistical power and higher presicion.

Although a high presicion is often comprehended as a sign of correct results, it is important to remember that with the size of data we have, spurios associations may also display high precision. Further, as we have been running multiple analyses in both of

the studies, so it is important to view the results together, and not to emphasise single results by the statistical significance.

Selection bias

There are different interpretations of selection bias. Conventionally, selection bias has been sees as a bias resulting from differing exposure-outcome associations between those included in a study and those eligible for the same study. Such bias may be caused by the selection of subjects or observations for inclusion, or factors influencing the study participation. A more stringent approach have been suggested where selection bias is seen as a consequence of conditioning on a common effect of the exposure and the outcome. (150) Within this framework, selection bias is clearly separated from confounding and also representativity. (151) Our use of population-based comprehensive registers with complete information and exact censoring should minimize selection bias to some degree. Although we have conditioned on being a patient for inclusion, as we do in both studies, we have assumed that our exposures were not associated with patient status. Thus, we do not think it is likely that patient status in an out-of-hours service is caused by any of the exposures in our studies.

For the follow up of the older patients in Papers II and III, there is a limitation from not having access to care registers with information on the use of nursing homes or home care. Patients admitted to nursing homes are still registered with their regular GP. However, they are mostly cared for by dedicated nursing home physicians who are not claiming reimbursement from Helfo. Thus these health services are not registered in the Control and Payment of Reimbursement Register. There is most likely a selection of only the frailest patients for nursing homes, which is highly related to their chance of admission to hospital, health service use, and death. However, with our within service design, we believe that this selection should be equally distributed between our exposures, reducing the possible differential loss to follow up for the elderly patients with the most severe conditions. (150)

For the study in Paper II, we only included claims with the physician identification numbers. This choice led to the exclusion of a large number of claims. If the patient contacts from claims with and without this physician identification number systematically differ, this could result in selection bias. Based on our knowledge of the data, we believe most of the claims without this number to come from out-of-hours physicians with fixed salaries (for out-of-hours work) from the municipalities. In Norway, such salary conditions are often offered for night shifts and, in some municipalities, also for day and late shifts in some municipalities. Out-of-hours physicians may differ in their activity, such as coding and referral rates, based on having fixed salaries or being self-employed. However, we believe that our matching, where we do not compare physicians between services, counteracts the potential effect from such selection, as most physicians within the same municipality are offered the same conditions for out-of-hours work within the same shifts. (137)

Information bias

Information bias is a skewness of the results caused by measurement errors in the data. (152) We usually distinguish between differential and non-differential measurement errors. Non-differential measurement errors are misclassification in exposure not systematically associated with the outcome, or misclassification of outcome not associated with the exposure. In our register data, this could be due to, for example, typing errors in the diagnose codes or time of contact. Non-differential measurement errors in the exposure most often give a smaller effect estimate (bias towards null). Also, there may be non-differential measurement error in the adjustment variables, leading to residual confounding. Differential measurement errors are systematically related to the values of the exposure or outcome and can cause spurious associations in any direction.

In register-based studies, we use data that is collected for other measures than research purposes. As the data from the Norwegian Patient Register and the Control and Payment of Reimbursement Register is commonly used as a base for funding, it is likely to be relatively accurate and complete. However, there have been cases of dishonest claims for reimbursement, and we cannot rule out the possibility of such claims

affecting our data. We believe such potentially false claims not to be related to the GP characteristics under study and thus to represent a non-differential measurement error. For Paper II, we made sensitivity analyses excluding the contacts from physicians with a very high number of claims (> 10,000 per 2 years, or 435 per month, representing the 98 percentile), as these could represent wrong registrations. Notably, this exclusion did not affect our results.

It is an advantage for researchers to know how data is generated. Our data is based on the registration of codes for diagnoses and procedures for claiming reimbursement. From my experience as a physician, I know that such registration tends to be less accurate with the increasing workload or acuity. Diagnosis coding in primary care (ICPC-2) is often a low priority when handling emergency cases. This is most likely why ICPC-2 codes from the out-of-hours services are shown to have low accuracy and validity.(40) Based on this knowledge, we chose not to use these codes when selecting diagnoses to include in our variable "critical conditions" in Paper II.

Our study design relies upon the variables time of contact and patient age to be quite accurate in the registrations. Such errors may have an even higher effect when we condition on time in our variables: Our main outcome variable in Paper II (unplanned hospital admission) and exposure variable in Paper III (referral) is based on the occurrence of registered contact in two registers within ten hours, namely a registered assessment by an out-of-hours physician (KUHR) followed by a registered unplanned contact with a hospital ward (NPR). We assume that such registration reflects an assessment by an out-of-hours physician leading to a referral to the hospital for most of the cases. However, this premise could be violated in (at least) two ways: A patient assessed by an out-of-hours physician and consequently referred to the hospital can choose for some reason not to go to the hospital as recommended. This will result in a wrong referral registration for the physician and give a smaller effect estimate. Further, a patient assessed by an out-of-hours physician and *not* referred but who deteriorate and need an acute admission within the following ten hours would be registered as admitted or referred by the physician from the index consultation. We believe that this was likely to be non-differential and not systematically related to the physician characteristic to

affect our results since registration in specialised health was gathered from another register out of the control of the out-of-hours physician.

Our estimation of the municipality where the out-of-hours service took place is another potential source of measurement error. The estimation method is described closely in the methods section and potentially lead to comparing activity from different services in the matched groups if the estimated municipalities are wrong. Out-of-hours services from smaller and cooperating municipalities with fewer contacts are more likely to be affected by these errors. If the matching is compromised, this could introduce residual confounding. However, our balance test of confounders counts against this.

For the other registers we have used, namely The Cause of death registry, Statistics Norway, and the regular General Practitioner register, we believe the data to have high accuracy for our use. However, the non-complete registration of GP locums in the regular General practitioner register, especially for the first years of the study period, may have limited the inclusion of claims made by GP locums in our study and thus limited our sample size.

Confounding

Confounding is a distortion of the association between an exposure and an outcome caused by a common cause of both the exposure and the outcome. (128) Confounding is a typical threat to observational analyses. Commonly, confounding is handled by including known confounding variables in regression analyses, referred to as statistical adjustment. However, such adjustments have strong assumptions that are seldom satisfied in health services registry data. For our studies, we assumed the exposure-outcome association likely confounded by both measurable and unmeasurable confounding factors; thus, we believed only adjusting for the measured confounders would leave us with biased results. Therefore, handling potential confounding was one of our main priorities when planning and designing our studies. Our measures to address confounding are closely described in the methods section.

In Paper II, our main concern was the potential systematic relation between the patients' characteristics, physician characteristics and patient outcomes. We handled this by closely matching observations from patients we assumed to be comparable. Hence, no other likely factors were causing specific physicians to see certain types of patients within our groups. The results from our sensitivity analyses, showing independence between the patient and physician characteristics, supported these assumptions. However, we can still not rule out confounding, which is not reflected in our measured variables.

In Paper III, confounding by indication was our main concern. This also arises from a common cause of the exposure and the outcome, in our case, the patient condition (confounder) affecting both the indication for being referred (exposure) and the indication for further health service use and risk of death (outcome). (131) We assumed it highly unlikely that we could handle this confounding by adjustments in the model. In Paper III, we addressed this by applying an instrumental variable design to estimate the causal effect of the exposure on the outcome. The process and assumptions for using this method are carefully described in the methods section. There are several potential limitations to applying this method, as it depends on strong assumptions both of the instrumental variable model and the matching design of the study.

For the study in Paper III, we were depending on information about the physicians that was strongly predictive of their referral preferences. This was achieved by using activity information from each physician's activity over up to 9 years in out-of-ours work. Using physician referral preference as an instrument was based on our assumptions about the variability in physicians' thresholds for referrals. However, as each physician's actual threshold for referral is not measurable, we had to find a measurable variable as a proxy. Thus, we chose to use the variation in referral rate as a proxy for the physician's referral threshold or preference. In Paper II, we based our variable physician admission proportion on the activity during the previous four months only. However, for our use in Paper III, we needed increased statistical power; thus, we chose to include all out-of-hours activity for each physician, for male and female

patients separately. Further, the matching of the observations was less strict than in Paper II, as the patients were more similar, due to including only elderly patients, who had not been assessed by the same physician earlier in the study period and were not frequent attenders the year of the contact. The matching was thus only based on combining observations from patients visiting the same out-of-hours service, shift and patient sex within each year. This enabled higher precision, which was necessary due to the power demanding instrumental variable analyses, and further that inclusion of only elderly patients also affected the study power.

The assumptions for expecting valid results from instrumental variable analyses are closely described in the methods section. We tested the relevance assumptions by estimating the instrument-exposure association (physicians´ referral preference and their subsequent referrals decisions). The results indicated strong associations, with F-values of ≈1,200. Although such a dichotomisation may be misleading, an F-test >10 is conventionally regarded as sufficient. (144) The independence assumption is not testable; however, our balance test of confounders showing weak or no associations between a set of potential confounders and the physician referral preference supports this assumption. We also assumed that the referral preference variable only affected the outcome via hospital referral for the index patient. If the effect on the outcome of being assessed by a physician with higher referral preference is mediated by another factor than the effect on the referral, this may violate the exclusion restriction and possible bias the effect estimate. A possible example may be if physicians with higher referral preferences also arrange for better treatment plans for the patients they do not refer. However, as time is scarce in the out-of-hours setting, most treatment actions could more likely be seen as alternatives to referral.

In addition to performing instrumental variable analyses, we chose to perform multivariable adjusted regression analyses, allowing us to assess the degree of confounding even after adjusting for all measured confounding. As we expected, the multivariable adjusted analyses showed a substantially increased risk of dying following a referral to hopital. Based on our confidence in the health services, we do not believe this to be valid evidence for recommending not to refer the patients. Also, the balance

test of suggested confounding variables showed an unequal distribution of these confounders in the groups. Hence, patients with previous health services use and higher age were more likely to be assessed by physicians with higher referral preference. This unequal distribution of confounders from the multivariable analyses supports our use of an instrumental variable approach.

An important limitation of the instrumental variable approach is the estimation of the risk of death. Several methods have been suggested in instrumental variable analyses for dichotomous outcomes; however, they all have their limitations. (144) Also, using these methods depends upon a scaling of the estimated effect according to the instruments-exposure association. In practice, this means that the effects of being assessed by physicians with different thresholds for referral on the risk of death are scaled to show the effect of being referred compared to not being referred. That is, the actual difference in mortality between patients visiting different physicians was small. Our estimations of substantially reduced mortality for patients who are referred based upon their physicians' referral preference depend on heavy assumptions and thus must be interpreted with caution. It is further essential to remember that this effect does not apply for all patients, but only for the patients who will be affected by, i.e. have their referral decision based on the physician threshold. This will be elaborated on further in the next sections.

External validity – Generalisability

The findings presented in Papers II and III are based on population-based comprehensive health registers, supporting the external validity. There is reason to believe that our findings can be generalised to other populations with similar health care systems, particularly with gatekeeping. However, our findings are likely valid only for the patient population using the out-of-hours services, known to have higher morbidity and conditions with higher acuity than the general population. For Paper III, the results are valid only for the patient population who will have their referral decision determined by the referral threshold of the physician assessing them. From a policy angle, this is also the most interesting group since they would be the target population of changing

the referral threshold in out-of-hours services. As described in the methods section, there is no perfect way of predicting which of the patients will be compliers; hence, for which of the patients the instrument will have an effect on the exposure. This is only based on empirical knowledge of the situations where physicians may choose systematically differently regarding referral. However, to try to estimate the variability of the instrument, we made analyses where we estimated the effect of one standard deviation change in the physician referral preference. These results showed that a change of one standard deviation in the physician referral threshold (instrument) was associated with a risk difference of about four percentage points for referral to the hospital with a F-value of about 1,200.

6.3 Discussions of findings

In the following section, I will discuss the relevance and the interpretation of the main findings in a more general context. For a more specific discussion of the results in each Paper, please see the discussion sections in each Paper.

The load of high-quality health services provided to people worldwide have probably never been higher, and the Norwegian health services are currently ranged as top of the class, based on the health outcome, administrative efficiency and access to care. (3, 153) However, both nationally and internationally, health services face major challenges with increasing population needs and expectations, and rising costs. If this trend continues, the demands will most likely exceed the current capacity of the health services; thus, there is an evident need for action. This thesis sheds light on the complexity of the patient dynamics and some aspects of the possible measures to handle these increasing demands.

Policymakers have acknowledged and addressed the challenges of increasing demands and costs by adjusting and reorganising the current systems. Specialised services have been given particular attention due to immense costs and long waiting lists. Consequently, tasks and responsibilities have been transferred to primary care to relieve specialised services. (6) The GPs in our study perceived workload also in primary care

as increasing, in line with other studies and reports showing an increasing number of consultations and administrative work reflected as a higher number of working hours. (32, 39, 109, 110) However, the GPs listed a range of various causes for the increased workload, not only related to the coordination reform. Apparently the causes differed markedly between the practices. These findings suggests that relieving the workload by targeting the specific causes likely requires a broad approach. However, a common feature for all our participants was the perception of too much work per GP. A possible solution would thus be to increase the number of personnel in general practice. A plan of recruiting more GPs was stated in the coordination reform; yet, the preliminary increase in GPs seems not to cover the increased needs in the populations. (7) There are ongoing pilot-studies of introducing primary care teams where other health care personnel like nurses may contribute to relieving the GP of workload. (113)

The GPs in our study expressed severe concerns about the sustainability of the regular GP scheme, and several of them had considered leaving their work, even if they had "the best job in the world". This finding aligns with the current challenges of recruitment and retention of GPs seen in Norway and England. Over the latest years, not only rural municipalities but also larger cities have struggled to recruit and retain GPs, leading to a likely increased workload for the GPs left in the scheme. The recruitment and retention problems are also suggested to directly affect the use of health care through affected continuity of care for the patients. (154, 155)

Further, the composition of GPs staffing the regular GP scheme may be affected. The mean age of the current regular GPs is relatively high and is closely related to their GP experience. Thus, recruiting new regular GPs will most likely reduce the workforce's mean age and years of experience. An increasing share of female Norwegian medical students, will likely affect the sex composition of future physicians. As the legislation now requires all regular GPs to enter education programs to achieve specialist status in general practice, a higher share of GPs will be specialists.

The Norwegian Medical Association and the media have outlined that a change in the composition towards younger GPs with ambitions of starting family life may not be

compatible with the working hours and responsibility required from being a regular GP, and urged a need for action (156-158) Also, the doctor's role has arguably changed; whereas older physicians regard it as a lifestyle, the younger physicians see it more as a job. (159) This results in potential work-home interface stress, also found to predict emotional exhaustion among physicians. (160) The GPs in our study discussed this work-home balance and pointed to their limited possibilities to participate in their own family life as reasons for perceiving the workload as too heavy. The participants without family responsibilities highlighted this as an advantage for coping. Thus, persistent heavy workload in general practice may serve as a selection mechanism affecting the GPs staffing the service, and affect their participation in the out-of-hours services. (35)

The GPs serve as the primary providers in the out-of-hours services, hence the composition of GPs will also likely be reflected in the composition of out-of-hours physicians. Out-of-hours services are a crucial part of the emergency medical chain and primary care gatekeeper function, responsible for 1/3 of unplanned admissions. (43) Thus, changes in the referral rates from out-of-hours services may substantially affect the inflow of patients to the hospitals. Our findings suggest that the characteristics of the GPs staffing the out-of-hours services affects gatekeeping. However, the expected change towards younger and more female GPs do not seem to have large effects on health care use, costs, or patient safety. Currently, about half of the out-of-hours services are staffed by GPs, half of these are specialists, and many have high experience both from daytime GP services and out-of-hours work. The change in the composition may reflect as GPs with less experience both from daytime and out-of-hours work possessing the important gatekeeping role. In line with less experienced physicians potentially being more risk averse, (84) our findings suggest that this may lead to a higher number of unplanned admissions, with emergency department crowding as a potential consequence.

We found the physician admission proportion seemingly being the strongest predictor of future admission practice. This may have several implications. Firstly, this demonstrates a variation in health services provision that is not related to the patients' health

condition. This may imply a potential for reducing costs by reducing this variation. Of the patients aged 70 years and older assessed by an out-of-hours physician in our studies, about one out of four had an immediate unplanned admission. If the assumptions of our model hold, there are almost two-fold odds for an elderly patient being admitted solely due to meeting a physician in the highest quarter compared to the lowest quarter of prior admission proportion. We found that this was costly for the services (Paper II), not surprising, as such admissions have an average of 3 days in hospital (Paper III). However, an important nuance to this was the association between higher admission proportion and admitting more patients with critical conditions (Paper II). This suggests that the observed variation in referral rate was not simply due to excess admissions from the high referrers, and supports caution in the use of observed variation for quality assessments in health services. Together, these results also support a need for enhancing the conditions for referral decisions, to increase the referral accuracy. The GPs in our interview study expressed that easily accessible hospital colleagues for conferring could have prevented many unplanned admissions when the GPs were in doubt. Making hotlines for conferring with an experienced hospital specialist is being piloted in some Norwegian hospitals. (50, 51)

Secondly, our findings support the theory that physicians have differing thresholds for admission and referrals. Variation in referral rates have been thoroughly studied, and both national and international studies have found up to five-fold differences in referral rates from primary care physicians. (10, 75, 77) Although a substantial part of this variation may be due to systematic differences in the patient population and organisational factors between the physicians, such factors do not seem to explain all the variation, thus suggesting personal factors of the physician as influential. This has been further supported by qualitative research, questionnaires and vignette studies, where personal factors like tolerance of uncertainty, fear of litigation and low professional confidence have been identified. (76, 81, 84, 161) In out-of-hours services, the referral decision is often made under stressful circumstances, with limited time, resources, and patient knowledge. Such circumstances are known to affect decision making in the direction of using simple heuristics (162) possibly making the physician more prone to act in the “usual way”, relying more on the personal threshold. Such

effects may also affect decision making and thus gatekeeping through a high workload over time. (163)

To change habits, and improve diagnostic accuracy, feedback is a key feature. (86, 164) As feedback on referral decisions are not systematised in the out-of-hours services, such learning effects can be hard to obtain. (76) This may be why out-of-hours GPs are shown to be poor at judging their own referral practices. (81) Even experienced physicians, with high clinical confidence are shown to have low clinical accuracy if they are not given feedback. (161, 165) Providing the out-of-hours physicians with more systematic feedback on patient outcomes could hopefully contribute to increasing diagnostic and referral accuracy and should be ensured. Privacy provisions and the diversity of computer systems have hindered such feedback; however, this may be enabled with the implementation of the new Health Platform.

Tolerating clinical uncertainty has been targeted as impacting physicians decision making. (166, 167) However, closely related is also the patients' tolerance of uncertainty, and trust in the health services. (168) The increasing general knowledge and access to specific information in the population are suggested to impact the trust in the health services. Earlier, the doctor's decision was often highly trusted. However, this trust is now being challenged, potentially impacting referral decisions. (1) Shared decision-making has been suggested to lower the use of specialised care. (66) However, for this to have the desired effect, time is crucial to allow for a reasonable discussion and explanation of the alternatives to escalated care. (169) The empowerment of the patients and next-of-kin may result in the physician accepting the demands of a referral, particularly in stressful situations. Our participants also brought up this potential mechanism and regarded it as a threat to the gatekeeper role. In the national and international campaigns "Choosing wisely", shared decision making is emphasised to reduce specialised care. (92) However, it is not known whether this is possible to implement in current the out-of-hours services, without changing the limited time frames of the out-of-hours consultations.

As the variation in referral rates have been scrutinized, the consequences of this variation have not previously received the deserved attention. There is reason to believe that in some cases, such as where patients present with diffuse symptoms and findings that may imply severe illness, some physicians will choose to refer, while others will choose a wait and see approach. In our third paper, we tried to approach the effects of a referral decision made in such cases, especially for patients relatively unknown to the physician and the out-of-hours system.

Assuming our model holds, our results suggest that if gatekeeping is impacted towards referring more patients, this results in a high workload on hospitals and specialised care, as patients with referrals attributable to their physicians' threshold receive substantial specialist care. Further such referrals do not seem to relieve the GPs, in contrast to what could be expected when specialised care "handle the patient's problem". This finding rather suggests that the coordination reform works as intended, with some tasks transferred to primary care. Thus, it is essential to ensure that the patients who do not need a referral do not get one, both to prevent unnecessary use of the limited capacity in specialised care and protect the patients from potentially harmful investigations and treatments. Especially elderly frail patients are vulnerable to hospital admissions. The current campaigns aimed at reducing medical overactivity mostly applies to elective care. Nevertheless, with the increasing demands and limited hospital capacity, the reduction of acute referrals will likely also be targeted.

Our findings further suggest that referring too few patients may also have negative consequences for patients affected by the physician threshold if they do not get the care they need. Our instrumental variable approach, scaling the effect of the differences in physician threshold onto the effect of referral or not for these patients, implies a substantially reduced mortality for the referred patients. This further implies that simply asking physicians to lower their referrals to relieve the specialised care of pressure may be detrimental to patient safety, an important aspect of the discussion on reducing the inflow of patients to specialised care. If such measures are to be applied, our findings suggest thorough evaluations before implementation, justifying the use of randomised controlled trials.

However, based on limited capacity in hospitals, our results also indicate the consequences of exceeding the capacity in the emergency departments resulting in an automatically stricter selection of the patients. There is reason to believe that such limited access will not affect patients with obvious needs of admissions, like a hip fracture, but rather the patients where the physicians doubt the necessity of the referral. According to our results, such situations may be very harmful to these patients.

The findings presented in this thesis outline the need for improving the framework for decision making in out-of-hours services to improve referral accuracy. High costs from specialised services associated with potentially unnecessary hospital admissions and the potential impact on patient safety from raising the referral threshold without improving accuracy will justify heavy measures to enhance gatekeeping conditions. Such measures may involve securing experienced physicians to staff the out-of-hours services, enabling easy access for conferral with colleagues in specialised services, or allowing for prolonged observation time of patients, where there is doubt about the referral. Obtaining this may be aided by further centralisation of the out-of-hours services. Also, securing the physicians' learning effects by giving them feedback on their referral decisions may be another important measure. Notably, based on the knowledge conveyed in this thesis, simply providing physicians with their referral rates for comparisons will not be sufficient. Without the financial incentives to enhance the conditions for gatekeeping, the investments required may be hard to prioritise, as there are so many other needs to be covered in primary care. However, the prize may be preserving what is currently ranged as the best health care system in the world.

6.4 Future research

This thesis sheds light on different aspects of the complexity of health services. Based on the methods used and our results, I will suggest implications for future research. Our demonstration of how changes in one part of the services may cause unintended 'spillover effects' in other parts of the service emphasise the need for a broad perspective, with outcome measures in different levels when planning future research.

The same applies when planning for quality assessment and evaluating measures for improvement in health services.

We have further demonstrated a potential use of register data to answer causal questions, with different approaches to address the possible limitations from such data carefully. However, there is a need for a thorough knowledge of both the data and reality reflected by the data to minimise pitfalls from systematic errors such as selection bias, information bias and confounding. We hope that establishing and making available registers for more parts of the service, like nursing homes and emergency medical services such as ambulance services, will serve to avoid “holes” in the patients’ trajectories. The municipal patient and user register (Kommunalt pasient og brukerregister, KPR) is already established, and there are plans to also make registers for the emergency medical services. (13) Access to such registers would allow for a more comprehensive approach to study complex patient dynamics and prevent bias from loss to follow up. Moreover, we have hopes for shortening the time span between the approval of the application and the availability of the data. In our case, this process took almost hel, which may lead to outdated data for some research purposes. (69)

Our results from Paper II suggest a potential association between out-of-hours physicians with low out-of-hours experience and increased risk of death for their patients; however, the confidence intervals were wide. This finding justifies further research, particularly as the expected changes in the composition of out-of-hours physicians may involve a higher share of inexperienced physicians. Finally, our findings from Paper III of potentially reduced mortality for patients referred to hospital in cases where the indications for referrals were not clear cut justifies further research such as the use of randomised controlled trials to assess the potential consequences of altering the referral threshold.

7 Conclusions

The overarching objective of this thesis was to explore the complex dynamics in the health services, emphasising the effects of the increasing demands on health services and how measures to reduce pressures in one part of the service may reflect as increasing pressure in other parts or even compromised patient safety. We have demonstrated how factors affecting the composition of physicians holding the important gatekeeper role may affect the inflow of patients to hospitals and the further use of health services. We also display the potential consequences for patient safety by simply asking these gatekeepers to raise their referral thresholds, underlining the need for a thorough evaluation of measures taken to reduce pressure in specialised health services. Collectively, our findings serve as knowledge for policymakers planning health services and also emphasise the need to prioritise primary care and the general practitioner service to preserve our high-quality health services.

8 References

1. Fritz Z, Holton R. Too much medicine: not enough trust? J Med Ethics. 2019;45(1):31-5.
2. Fullman N, Yearwood J, Abay SM, Murray CJL, Lozano R. Measuring performance on the Healthcare Access and Quality Index for 195 countries and territories and selected subnational locations: a systematic analysis from the Global Burden of Disease Study 2016. The Lancet. 2018;391(10136):2236-71.
3. OECD. Health at a glance 2021 [Available from: https://www.oecd.org/health/health-at-a-glance/.
4. Saunes IS, Karanikolos M, Sagan A. Norway: Health system review. Health systems in Transition, 2020. World Health Organization. Regional Office for Europe, European Observatory on Health Systems and Policies 2020.
5. The Norwegian Directorate of Health. Status and development for specialised health services (Status og utviklingstrekk for spesialisthelsetjenesten. Kostnader, aktivitet og kvalitet. Samdata spesialisthelsetjeneste 2019). 2019. Report No.: IS-2957.
6. The Ministry of Health and Care Services (Helse-og omsorgsdepartementet). The Coordination reform (Meld. St. 47 (2008–2009). Samhandlingsreformen – Rett behandling – på rett sted – til rett tid.). 2009.
7. The Norwegian Directorate of Health. Oppfølging av Handlingsplan for allmennlegetjenesten 2020-2024. 2020 19 May.
8. Act relating to specialist health services (Lov om spesialisthelsetjenesten m.m. (spesialisthelsetjenesteloven)), (2001).
9. The Norwegian Directorate of Health. The referral guideline (Henvisningsveilederen). In: Health TNDo, editor. 2015.
10. Ringberg U, Fleten N, Forde OH. Examining the variation in GPs' referral practice: a cross-sectional study of GPs' reasons for referral. The British journal of general practice : the journal of the Royal College of General Practitioners. 2014;64(624):e426-33.
11. McBride D, Hardoon S, Walters K, Gilmour S, Raine R. Explaining variation in referral from primary to secondary care: cohort study. BMJ. 2010;341:c6267.
12. Lasserson D, Smith H, Garland S, Hunt H, Hayward G. Variation in referral rates to emergency departments and inpatient services from a GP out of hours service and the potential impact of alternative staffing models. Emerg Med J. 2021.
13. The Norwegian Directorate of Health. The Nordic Emergency Medical Services Benchmarking Report 2014-2018.; 2018.
14. Act of Patient Rights (2001) (Pasient og brukerrettighetsloven) (2001).
15. Laugesen K, Ludvigsson JF, Schmidt M, Gissler M, Valdimarsdottir UA, Lunde A, et al. Nordic Health Registry-Based Research: A Review of Health Care Systems and Key Registries. Clin Epidemiol. 2021;13:533-54.
16. Ambugo EA, Hagen TP. Effects of introducing a fee for inpatient overstays on the rate of death and readmissions across municipalities in Norway. Soc Sci Med. 2019;230:309-17.
17. Norwegian health net (Norsk helsenett). Out of-pocket spending for health services (Eigendelar for helsetenester): Helsenorge; 2021 [Available from:

https://www.helsenorge.no/betaling-for-helsetjenester/egenandeler-for-helsetjenester/.
18. Eide TB, Straand J, Bjorkelund C, Kosunen E, Thorgeirsson O, Vedsted P, et al. Differences in medical services in Nordic general practice: a comparative survey from the QUALICOPC study. Scand J Prim Health Care. 2017;35(2):153-61.
19. The Commonwealth Fund. International Health Care System Profiles, England 2013 [Available from: https://www.commonwealthfund.org/international-health-policy-center/countries/england.
20. Starfield B, Shi L, Macinko J. Contribution of primary care to health systems and health. Milbank Q. 2005;83(3):457-502.
21. Carlsen B, Norheim OF. Introduction of the patient-list system in general practice. Changes in Norwegian physicians' perception of their gatekeeper role. Scand J Prim Health Care. 2003;21(4):209-13.
22. The Norwegian Directorate of Health. Regular general practitioner statistics (Fastlegestatistikk) 2020 [August 23, 2021]. Available from: https://www.helsedirektoratet.no/statistikk/fastlegestatistikk.
23. Åmotsbakken MD. The regular general practitioners are ever more popular (Fastlegene stadig mer populære i befolkningen). 2017.
24. Statistics Norway. More regular GP and out-of-hours consultations than last year (Flere konsultasjoner hos fastleger og legevakt i fjor) 2020 [Available from: https://www.ssb.no/helse/artikler-og-publikasjoner/flere-konsultasjoner-hos-fastleger-og-legevakt-i-fjor.
25. The Norwegian Medical Association. Information about the new specialisation in general practice (Informasjon om ny spesialisering i allmennmedisin) 2020 [Available from: https://www.legeforeningen.no/fag/spesialiteter/Allmennmedisin/informasjon-om-ny-spesialistutdanning-i-allmennmedisin/.
26. Regulation relating to a Municipal Regular GP Scheme (Forskrift om fastlegeordning i kommunene), (2012).
27. Helsedirektoratet. Fastlegestatistikk 2017 www.helsedirektoratet.no2018 [Available from: https://helsedirektoratet.no/statistikk-og-analyse/fastlegestatistikk#fastlegestatistikk-2017.
28. Statistics Norway. Good results in school, small change in choice of profession (Gode skoleresultater – liten endring i yrkesvalg) 2020 [Available from: https://www.ssb.no/utdanning/artikler-og-publikasjoner/gode-skoleresultater-liten-endring-i-yrkesvalg.
29. The Norwegian Public Health Act (Folkehelseloven) (2012).
30. The Act of Municipal Health and Care Services (Helse- og omsorgstjenesteloven) (2012).
31. Monstad K, Kaarbøe O, Holmås TH, Askildsen JE. Evaluation of municipal co-financing (Evaluering av kommunal medfinansiering). Tidsskrift for omsorgsforskning. 2016;2(02):135-41.
32. Morken T, Rebnord IK, Maartmann-Moe K, Hunskaar S. Workload in Norwegian general practice 2018 - an observational study. BMC Health Serv Res. 2019;19(1):434.
33. Verdens Gang. Kampen om fastlegene 2020 [Available from: https://www.vg.no/spesial/2017/fastleger/.

34. Trønderopprøret. Fastlegeordningen 2.0 Trønderopprøret 2019 [Available from: http://www.flo20.no/.
35. Iversen Tor. Mapping the current situation in the out-of-hours services (Kartlegging av situasjonen i dagens legevaktsordning). 2021.
36. Groenewegen P, Heinemann S, Gress S, Schafer W. Primary care practice composition in 34 countries. Health Policy. 2015;119(12):1576-83.
37. The Norwegian Medical Association (Legeforeningen). The correct diagnosis at the right time (Riktig diagnose til rett tid). 2017.
38. Forskrift om krav til og organisering av kommunal legevaktordning, ambulansetjeneste, medisinsk nødmeldetjeneste mv. (akuttmedisinforskriften), (2015).
39. Rebnord IK, Morken T, Maartmann-Moe K, Hunskaar S. Out-of-hours workload among Norwegian general practitioners - an observational study. BMC Health Serv Res. 2020;20(1):944.
40. Sandvik HH, S. Blinkenberg, J. . Yearly statistics from out-of-hours services (Årsstatistikk fra legevakt 2019). 2020.
41. Sandvik H, Hunskar S. [Working style among regular general practitioners and other doctors in the out-of-hours services]. Tidsskr Nor Laegeforen. 2010;130(2):135-8.
42. Steeman L, Uijen M, Plat E, Huibers L, Smits M, Giesen P. Out-of-hours primary care in 26 European countries: an overview of organizational models. Fam Pract. 2020;37(6):744-50.
43. Blinkenberg J, Pahlavanyali S, Hetlevik Ø, Sandvik H, Hunskaar S. General practitioners' and out-of-hours doctors' role as gatekeeper in emergency admissions to somatic hospitals in Norway: registry-based observational study. BMC Health Services Research. 2019;19(1).
44. Thorsen O, Hartveit M, Kristoffersen JE, Ringberg U. [The art of referring]. Tidsskr Nor Laegeforen. 2018;138(12).
45. Forskrift om prioritering av helsetjenester, rett til nødvendig helsehjelp fra spesialisthelsetjenesten, rett til behandling i utlandet og om klagenemnd (prioriteringsforskriften), (2001).
46. The Norwegian Directorate of Health. National recommendations, advice and patient pathways (Nasjonale anbefalinger, råd og pakkeforløp) 2021 [Available from: https://www.helsedirektoratet.no/produkter?tema=prioriteringsveileder.
47. Akbari A, Mayhew A, Al-Alawi MA, Grimshaw J, Winkens R, Glidewell E, et al. Interventions to improve outpatient referrals from primary care to secondary care. Cochrane Database Syst Rev. 2008(4):CD005471.
48. Eskeland SL, Rueegg CS, Brunborg C, Aabakken L, de Lange T. Electronic checklists improve referral letters in gastroenterology: a randomized vignette survey. Int J Qual Health Care. 2018;30(6):450-6.
49. Ree AO. Medisinsk-faglig innhold i henvisninger "Den gode henvisning». Helsedirektoratet; 2003. Contract No.: HIS 80322:2003KITH 22/03.
50. Sykehuset Østfold. Reccomendations on Cooperation (Anbefalinger om samhandling) 2021 [Available from: https://01.fastlegeportalen.no/pko-nytt/suff/anbefalinger-om-samarbeid/.
51. Fagutvalg leger. Recommendations about routines for cooperation between regular general practitionaers and St. Olavs hospital (Anbefalinger om

samarbeidsrutiner mellom fastleger og St. Olavs hospital). Helse Midt Norge; 2019. Contract No.: Versjon 5.1 - 17.06.19.
52. Forrest CB. Primary care in the United States: primary care gatekeeping and referrals: effective filter or failed experiment? BMJ. 2003;326(7391):692-5.
53. Reibling N, Wendt C. Gatekeeping and provider choice in OECD healthcare systems. Current Sociology. 2012;60(4):489-505.
54. Sripa P, Hayhoe B, Garg P, Majeed A, Greenfield G. Impact of GP gatekeeping on quality of care, and health outcomes, use, and expenditure: a systematic review. The British journal of general practice : the journal of the Royal College of General Practitioners. 2019;69(682):e294-e303.
55. Velasco Garrido M, Zentner A, Busse R. The effects of gatekeeping: a systematic review of the literature. Scand J Prim Health Care. 2011;29(1):28-38.
56. Franks P, Clancy CM, Nutting PA. Gatekeeping revisited--protecting patients from overtreatment. N Engl J Med. 1992;327(6):424-9.
57. Lilford R, Mohammed MA, Spiegelhalter D, Thomson R. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. The Lancet. 2004;363(9415):1147-54.
58. Love T, Dowell AC, Salmond C, Crampton P. Quality indicators and variation in primary care: modelling GP referral patterns. Fam Pract. 2004;21(2):160-5.
59. Sullivan CO, Omar RZ, Ambler G, Majeed A. Case-mix and variation in specialist referrals in general practice. The British journal of general practice : the journal of the Royal College of General Practitioners. 2005;55(516):529-33.
60. Nicholl J. Case-mix adjustment in non-randomised observational evaluations: the constant risk fallacy. J Epidemiol Community Health. 2007;61(11):1010-3.
61. Norwegian Ministry of Health and Care Services. National Health and Hospital Plan 2020–2023

. In: Norwegian Ministry of Health and Care Services, editor. 2020.
62. McCarthy TW, K.L. The origins of health services research Health Services Research. 2000 Jun; 35(2): :375–87.
63. Skålevåg SA. Ignaz Semmelweis i Store norske leksikon på snl.no. Store Norske Leksikon2020.
64. Simensen AS. Florence Nightingale i Store norske leksikon på snl.no. 2020.
65. Donabedian A. Evaluating the quality of medical care. 1966. Milbank Q. 2005;83(4):691-729.
66. Wennberg JE. Unwarranted variations in healthcare delivery: implications for academic medical centres. BMJ. 2002;325(7370):961-4.
67. Greenland S, Morgenstern H. Confounding in health research. Annu Rev Public Health. 2001;22:189-212.
68. Rubin DB. For objective causal inference, design trumps analysis. The Annals of Applied Statistics. 2008;2(3).
69. Neuhauser D, Provost L, Bergman B. The meaning of variation to healthcare managers, clinical and health-services researchers, and individual patients. BMJ Qual Saf. 2011;20 Suppl 1:i36-40.
70. Centre for Clinical Documentation and Evaluation. Centre for Clinical Documentation and Evaluation (Senter for klinisk dokumentasjon og evaluering, SKDE) 2021 [Available from: https://helse-nord.no/skde.

71. Office of the Auditor General of Norway (Riksrevisjonen). Riksrevisjonens undersøkelse av myndighetenes arbeid med å sikre god henvisningspraksis fra fastlegene til spesialisthelsetjenesten 2018.
72. Barnett ML, Song Z, Landon BE. Trends in physician referrals in the United States, 1999-2009. Arch Intern Med. 2012;172(2):163-70.
73. Barnett ML, Bitton A, Souza J, Landon BE. Trends in Outpatient Care for Medicare Beneficiaries and Implications for Primary Care, 2000 to 2019. Ann Intern Med. 2021.
74. Ringberg U, Fleten N, Deraas TS, Hasvold T, Forde O. High referral rates to secondary care by general practitioners in Norway are associated with GPs' gender and specialist qualifications in family medicine, a study of 4350 consultations. BMC Health Serv Res. 2013;13:147.
75. Rossdale M, Kemple T, Payne S, Calnan M, Greenwood R. An observational study of variation in GPs' out-of-hours emergency referrals. British Journal of General Practice. 2007;57(535):152-4.
76. Calnan M, Payne S, Kemple T, Rossdale M, Ingram J. A qualitative study exploring variations in GPs' out-of-hours referrals to hospital. The British journal of general practice : the journal of the Royal College of General Practitioners. 2007;57(542):706-13.
77. O'Donnell CA. Variation in GP referral rates: what can we learn from the literature? Family Practice. 2000;17(6):462-71.
78. Thorsen O, Hartveit M, Johannessen JO, Fosse L, Eide GE, Schulz J, et al. Typologies in GPs' referral practice. BMC Fam Pract. 2016;17:76.
79. Bankart MJG BR, Rashid A, et al. Characteristics of general practices associated with emergency admission rates to hospital: a cross-sectional study. Emergency Medicine Journal 2011;2011(28):558-63.
80. Franks P, Williams GC, Zwanziger J, Mooney C, Sorbero M. Why do physicians vary so widely in their referral rates? J Gen Intern Med. 2000;15(3):163-8.
81. Ingram JC, Calnan MW, Greenwood RJ, Kemple T, Payne S, Rossdale M. Risk taking in general practice: GP out-of-hours referrals to hospital. The British journal of general practice : the journal of the Royal College of General Practitioners. 2009;59(558):e16-24.
82. Gehring ND, Kebbe M, Rathwell S, Perez A, Peng C, Zendher E, et al. Physician-related predictors of referral for multidisciplinary paediatric obesity management: a population-based study. Fam Pract. 2021.
83. Kristoffersen JE. Out-of-hours primary care and the patients who die. A survey of deaths after contact with a suburban primary care out-of-hours service. Scand J Prim Health Care. 2000;18(3):139-42.
84. Lawton R, Robinson O, Harrison R, Mason S, Conner M, Wilson B. Are more experienced clinicians better able to tolerate uncertainty and manage risks? A vignette study of doctors in three NHS emergency departments in England. BMJ Qual Saf. 2019;28(5):382-8.
85. Burton CD, McLernon DJ, Lee AJ, Murchie P. Distinguishing variation in referral accuracy from referral threshold: analysis of a national dataset of referrals for suspected cancer. BMJ open. 2017;7(8):e016439.
86. Omron R, Kotwal S, Garibaldi BT, Newman-Toker DE. The Diagnostic Performance Feedback "Calibration Gap": Why Clinical Experience Alone Is

Not Enough to Prevent Serious Diagnostic Errors. AEM Educ Train. 2018;2(4):339-42.
87. Nederhand ML, Tabbers HK, Splinter TAW, Rikers RMJP. The Effect of Performance Standards and Medical Experience on Diagnostic Calibration Accuracy. Health Professions Education. 2018;4(4):300-7.
88. Nassery N, Segal JB, Chang E, Bridges JF. Systematic overuse of healthcare services: a conceptual model. Appl Health Econ Health Policy. 2015;13(1):1-6.
89. Pathirana T, Clark J, Moynihan R. Mapping the drivers of overdiagnosis to potential solutions. BMJ. 2017;358:j3879.
90. National Health Services. Evidence-Based Interventions Programme 2021 [Available from: https://www.england.nhs.uk/evidence-based-interventions/.
91. The National System for Managed Introduction of New Health Technologies within the Specialist Health Service in Norway. The reconsideration project (Revurderingsprosjektet) 2021 [Available from: https://nyemetoder.no/om-systemet/revurdering-og-utfasing.
92. ABIM Foundation. Choosing Wisely 2021 [Available from: https://www.choosingwisely.org/.
93. Forster AJ, Asmis TR, Clark HD, Al Saied G, Code CC, Caughey SC, et al. Ottawa Hospital Patient Safety Study: incidence and timing of adverse events in patients admitted to a Canadian teaching hospital. CMAJ. 2004;170(8):1235-40.
94. Sager MA. Functional Outcomes of Acute Medical Illness and Hospitalization in Older Persons. Archives of Internal Medicine. 1996;156(6).
95. Perez T, Moriarty F, Wallace E, McDowell R, Redmond P, Fahey T. Prevalence of potentially inappropriate prescribing in older people in primary care and its association with hospital admission: longitudinal study. BMJ. 2018;363:k4524.
96. Tammes P, Purdy S, Salisbury C, MacKichan F, Lasserson D, Morris RW. Continuity of Primary Care and Emergency Hospital Admissions Among Older Patients in England. Ann Fam Med. 2017;15(6):515-22.
97. Huntley A, Lasserson D, Wye L, Morris R, Checkland K, England H, et al. Which features of primary care affect unscheduled secondary care use? A systematic review. BMJ open. 2014;4(5):e004746.
98. Gunther S, Taub N, Rogers S, Baker R. What aspects of primary care predict emergency admission rates? A cross sectional study. BMC Health Serv Res. 2013;13:11.
99. Posocco A, Scapinello MP, De Ronch I, Castrogiovanni F, Lollo G, Sergi G, et al. Role of out of hours primary care service in limiting inappropriate access to emergency department. Internal and Emergency Medicine. 2017;13(4):549-55.
100. Baird Bea. Understanding pressures in General Practice, The King´s fund. 2016
101. Groenewegen PP, Bosmans MWG, Boerma WGW, Spreeuwenberg P. The primary care workforce in Europe: a cross-sectional international comparison of rural and urban areas and changes between 1993 and 2011. Eur J Public Health. 2020;30(Suppl_4):iv12-iv7.
102. Thompson M, Walter F. Increases in general practice workload in England. The Lancet. 2016;387(10035):2270-2.
103. British Medical Association. Pressures in general practice 2021 [Available from: https://www.bma.org.uk/advice-and-support/nhs-delivery-and-workforce/pressures/pressures-in-general-practice.

104. Marchand C, Peckham S. Addressing the crisis of GP recruitment and retention: a systematic review. The British journal of general practice : the journal of the Royal College of General Practitioners. 2017;67(657):e227-e37.
105. Croxson CH, Ashdown HF, Hobbs FR. GPs' perceptions of workload in England: a qualitative interview study. The British journal of general practice : the journal of the Royal College of General Practitioners. 2017;67(655):e138-e47.
106. Hobbs FDR, Bankhead C, Mukhtar T, Stevens S, Perera-Salazar R, Holt T, et al. Clinical workload in UK primary care: a retrospective analysis of 100 million consultations in England, 2007–14. The Lancet. 2016;387(10035):2323-30.
107. National Health Service. A&E attendances and emergency admissions. 2021 [Available from: https://www.england.nhs.uk/statistics/statistical-work-areas/ae-waiting-times-and-activity/.
108. Hall LH, Johnson J, Watt I, Tsipa A, O'Connor DB. Healthcare Staff Wellbeing, Burnout, and Patient Safety: A Systematic Review. PLoS One. 2016;11(7):e0159015.
109. Norway S. Kortere pasientlister, lengre arbeidsdager? (Shorter patient lists, longer working days?) 2018 [Available from: https://www.ssb.no/helse/artikler-og-publikasjoner/kortere-pasientlister-lengre-arbeidsdager.
110. Rebnord IKE, Ole Johan; Hunskår, Steinar; Morken, Tone. The Regular GP's time use (Fastlegers tidsbruk). Uni Research Helse; 2018.
111. The ministry of Health and Care Services. Små kommuner har størst vaktbelastning: regjeringen.no; 2021 [Available from: https://www.regjeringen.no/no/aktuelt/sma-kommuner-har-storst-legevaktbelastning/id2891582/?utm_source=regjeringen.no&utm_medium=email&utm_campaign=nyhetsvarsel20211209.
112. The ministry of Health and Care Services. Handlingsplan for allmennlegetjenesten 2020-2024. 2019.
113. The Norwegian Directorate of Health. Primary Health Care Teams - Pilot project (Primærhelseteam - pilotprosjekt) 2018 [Available from: https://www.helsedirektoratet.no/om-oss/forsoksordninger-og-prosjekter/primaerhelseteam-pilotprosjekt.
114. Morris ZS, Boyle A, Beniuk K, Robinson S. Emergency department crowding: towards an agenda for evidence-based intervention. Emerg Med J. 2012;29(6):460-6.
115. Wyatt S, Child K, Hood A, Cooke M, Mohammed MA. Changes in admission thresholds in English emergency departments. Emerg Med J. 2017;34(12):773-9.
116. Salway RJ, Valenzuela R, Shoenberger JM, Mallon WK, Viccellio A. Emergency Department (Ed) Overcrowding: Evidence-Based Answers to Frequently Asked Questions. Revista Médica Clínica Las Condes. 2017;28(2):213-9.
117. Morley C, Unwin M, Peterson GM, Stankovich J, Kinsman L. Emergency department crowding: A systematic review of causes, consequences and solutions. PLoS One. 2018;13(8):e0203316.
118. Sara Marie Nilsen. Health care services under pressure – Consequences for patient flows, efficiency and patient safety in Norway 2020 [Available from:

https://www.researchgate.net/project/Health-care-services-under-pressure-Consequences-for-patient-flows-efficiency-and-patient-safety-in-Norway.
119. Pope C, Mays N. Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. BMJ. 1995;311(6996):42-5.
120. Malterud Kirsti. Kvalitative forskningsmetoder for medisin og helsefag: en innføring: Universitetsforlaget; 2017.
121. Malterud K. Systematic text condensation: a strategy for qualitative analysis. Scand J Public Health. 2012;40(8):795-805.
122. The Norwegian Directorate of Health. Control and Payment of Health Reimbursement Register Data base 2019 (KUHR-databasen 2019) 2019 [Available from: https://www.helsedirektoratet.no/tema/statistikk-registre-og-rapporter/helsedata-og-helseregistre/kuhr.
123. Bakken IJ, Ariansen AMS, Knudsen GP, Johansen KI, Vollset SE. The Norwegian Patient Registry and the Norwegian Registry for Primary Health Care: Research potential of two nationwide health-care registries. Scand J Public Health. 2020;48(1):49-55.
124. Norwegian Insitute of Public Health. The Norwegian Cause of Death Registry 2020 [Available from: https://www.fhi.no/en/hn/health-registries/cause-of-death-registry/.
125. Statistics Norway. Statistisk sentralbyrå 2021 [Available from: https://www.ssb.no/en.
126. The Norwegian Directorate of eHealth. The regular general practitioner register (Fastlegeregisteret) 2020 [Available from: https://helsedata.no/forvaltere/helsedirektoratet/fastlegeregisteret/.
127. Hamaker EL, Mulder JD, van IMH. Description, prediction and causation: Methodological challenges of studying child and adolescent development. Dev Cogn Neurosci. 2020;46:100867.
128. Hernán MA RJ. Causal Inference: What If: Boca Raton: Chapman & Hall/CRC; 2020.
129. Godøy A HI. Regional variation in healthcare utilization and mortality. 2018. Contract No.: No. 890.
130. Hernan MA, Hernandez-Diaz S, Werler MM, Mitchell AA. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. Am J Epidemiol. 2002;155(2):176-84.
131. Joseph KS, Mehrabadi A, Lisonkova S. Confounding by Indication and Related Concepts. Current Epidemiology Reports. 2014;1(1):1-8.
132. Hernan MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? Epidemiology. 2006;17(4):360-72.
133. Davies NM, Thomas KH, Taylor AE, Taylor GMJ, Martin RM, Munafo MR, et al. How to compare instrumental variable and conventional regression analyses using negative controls and bias plots. Int J Epidemiol. 2017;46(6):2067-77.
134. Lousdal ML. An introduction to instrumental variable assumptions, validation and estimation. Emerg Themes Epidemiol. 2018;15:1.
135. Swanson SA, Hernan MA. Commentary: how to report instrumental variable analyses (suggestions welcome). Epidemiology. 2013;24(3):370-4.

136. Angrist JD, Imbens GW, Rubin DB. Identification of Causal Effects Using Instrumental Variables. Journal of the American Statistical Association. 1996;91(434):444-55.
137. Goddager I. Fixed salaries in out-of-hours-services (Fastlønn legevakt) 2015 [Available from: https://www.regjeringen.no/contentassets/477c27aa89d645e09ece350eaf93fedf/no/sved/04.pdf.
138. OECD. OECD Data Elderly population 2021 [Available from: https://data.oecd.org/pop/elderly-population.htm.
139. Sandvik H, Hunskaar S. Frequent attenders at primary care out-of-hours services: a registry-based observational study in Norway. BMC Health Serv Res. 2018;18(1):492.
140. The Norwegian Directorate of Health. Activity based funding 2016 (Innsatsstyrt finansiering 2016). 2015. Report No.: ISBN-nr. 978-82-8081-417-3.
141. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. Journal of Chronic Diseases. 1987;40(5):373-83.
142. Correia S. A Feasible Estimator for Linear Models with Multi-Way Fixed Effects* 2016 [Available from: http://scorreia.com/research/hdfe.pdf.
143. Inoue A, Solon G. Two-Sample Instrumental Variables Estimators. Review of Economics and Statistics. 2010;92(3):557-61.
144. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. Stat Methods Med Res. 2017;26(5):2333-55.
145. Burgess S, Thompson SG, Collaboration CCG. Avoiding bias from weak instruments in Mendelian randomization studies. Int J Epidemiol. 2011;40(3):755-64.
146. van den Hoonaard WC. Is Anonymity an Artifact in Ethnographic Research? Journal of Academic Ethics. 2003;1(2):141-51.
147. Carlsen B, Glenton C. What about N? A methodological study of sample-size reporting in focus group studies. BMC Med Res Methodol. 2011;11:26.
148. Rothman KJ GS, Lash TL. Modern Epidemiology Wilkins LW, editor. Philadephia2008.
149. Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. Eur J Epidemiol. 2016;31(4):337-50.
150. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. Epidemiology. 2004;15(5):615-25.
151. Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. Int J Epidemiol. 2013;42(4):1012-4.
152. Rothman K. Epidemiology: an introduction. 2nd edition ed. . 2nd edition ed. Press OU, editor. New York2012.
153. Eric C. Schneider AS, Michelle M. Doty, Roosa Tikkanen, Katharine Fields, Reginald D. Williams II. Mirror, Mirror 2021: Reflecting Poorly: The Commonwealth Fund; 2021 [Available from: https://www.commonwealthfund.org/publications/fund-reports/2021/aug/mirror-mirror-2021-reflecting-poorly.

154. Sandvik H, Hetlevik O, Blinkenberg J, Hunskaar S. Continuity in general practice as predictor of mortality, acute hospitalisation, and use of out-of-hours care: a registry-based observational study in Norway. The British journal of general practice : the journal of the Royal College of General Practitioners. 2021.
155. Skarshaug LJ, Kaspersen SL, Bjorngaard JH, Pape K. How does general practitioner discontinuity affect healthcare utilisation? An observational cohort study of 2.4 million Norwegians 2007-2017. BMJ open. 2021;11(2):e042391.
156. Brekke M, Bjorke JS, Wisborg T. Where did all the doctors go? Tidsskr Nor Laegeforen. 2021;141(16).
157. The Norwegian Medical Association. Hovedbudskap som kan brukes i intervjuer eller kronikker 2020 [Available from: https://www.legeforeningen.no/contentassets/a740434ceb2649489cac566682c8ce98/hovedbudskap-som-kan-brukes-i-intervjuer-eller-kronikker.pdf.
158. Forening YL. - Man kan være en dyktig og seriøs lege selv om man ønsker å hente i barnehagen YngreLeger.no2019 [Available from: https://yngreleger.no/artikkel/man-kan-v%C3%A6re-en-dyktig-og-seri%C3%B8s-lege-selv-om-man-%C3%B8nsker-%C3%A5-hente-i-barnehagen.
159. Hertzberg TK, Skirbekk H, Tyssen R, Aasland OG, Ro KI. The good doctor - strong and persevering. Tidsskr Nor Laegeforen. 2016;136(19):1631-4.
160. Hertzberg TK, Ro KI, Vaglum PJ, Moum T, Rovik JO, Gude T, et al. Work-home interface stress: an important predictor of emotional exhaustion 15 years into a medical career. Ind Health. 2016;54(2):139-48.
161. Meyer AN, Payne VL, Meeks DW, Rao R, Singh H. Physicians' diagnostic accuracy, confidence, and resource requests: a vignette study. JAMA Intern Med. 2013;173(21):1952-8.
162. Kahneman D. Thinking, fast and slow Farrar, Straus and Giroux.; 2011.
163. Kazemi R, Haidarimoghadam R, Motamedzadeh M, Golmohamadi R, Soltanian A, Zoghipaydar MR. Effects of Shift Work on Cognitive Performance, Sleep Quality, and Sleepiness among Petrochemical Control Room Operators. J Circadian Rhythms. 2016;14:1.
164. Schiff GD. Minimizing diagnostic error: the importance of follow-up and feedback. Am J Med. 2008;121(5 Suppl):S38-42.
165. Meyer AND, Singh H. Calibrating how doctors think and seek information to minimise errors in diagnosis. BMJ Qual Saf. 2017;26(6):436-8.
166. Ilgen JS, Eva KW, de Bruin A, Cook DA, Regehr G. Comfort with uncertainty: reframing our conceptions of how clinicians navigate complex clinical situations. Adv Health Sci Educ Theory Pract. 2019;24(4):797-809.
167. Zwaan L, Hautz WE. Bridging the gap between uncertainty, confidence and diagnostic accuracy: calibration is key. BMJ Qual Saf. 2019;28(5):352-5.
168. Simpkin AL, Schwartzstein RM. Tolerating Uncertainty - The Next Medical Revolution? N Engl J Med. 2016;375(18):1713-5.
169. Yahanda AT, Mozersky J. What's the Role of Time in Shared Decision Making? AMA J Ethics. 2020;22(5):E416-22.

Paper I

Svedahl et al. BMC Family Practice (2019) 20:68
https://doi.org/10.1186/s12875-019-0952-5

BMC Family Practice

RESEARCH ARTICLE Open Access

Increasing workload in Norwegian general practice – a qualitative study

Ellen Rabben Svedahl[1*], Kristine Pape[1], Marlen Toch-Marquardt[1], Lena Janita Skarshaug[1], Silje-Lill Kaspersen[1,2], Johan Håkon Bjørngaard[1,3] and Bjarne Austad[1]

Abstract

Background: General practitioners (GPs) play a key role in securing and coordinating appropriate use of healthcare services, by providing primary and preventive healthcare and by acting as gatekeepers for secondary healthcare services. Historically, European GPs have reported high job satisfaction, attributed to high autonomy and good compatibility with family life. However, a trend of increasing workload in general practice has been seen in several European countries, including Norway, leading to recruitment problems and concerns about the well-being of both GPs and patients. This qualitative interview study with GPs and their co-workers aims to explore how they perceive and tackle their workload, and their experiences and reflections regarding explanations for and consequences of increased workload in Norwegian general practice.

Methods: We conducted seven focus groups and four individual interviews with GPs and their co-workers in seven GPs' offices in Mid-Norway: three in rural locations and four in urban locations. Our study population consisted of 21 female and 12 male participants; 23 were GPs and 10 were co-workers. The interviews were analysed using systematic text condensation.

Results: The analysis identified three main themes: (1) Heavy and increasing workload – more trend than fluctuation?; (2) Explanations for high workload; (3) Consequences of high workload. Our findings show that both GPs and their co-workers experience heavy and increasing workload. The suggested explanations varied considerably among the GPs, but the most commonly cited reasons were legislative changes, increased bureaucracy related to documentation and management of a practice, and changes in patients' expectations and help-seeking behaviour. Potential consequences were also perceived as varying, especially regarding consequences for patients and the healthcare system. The participants expressed concerns for the future, particularly in regards to GPs' health and motivation, as well as the recruitment of new GPs.

Conclusions: This study found heavy and increasing workload in general practice in Norway. The explanations appear to be multi-faceted and many are difficult to reverse. The GPs expressed worries that they will not be able to provide the population with the expected care and services in the future.

Keywords: General practitioner, GP, General practice, Family medicine, Workload, Coordination reform, Norway, Qualitative, Interview study

* Correspondence: ellen.r.svedahl@ntnu.no
[1]Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Håkon Jarls gate 11, 7030 Trondheim, Norway
Full list of author information is available at the end of the article

Introduction

General practitioners (GPs) play a key role in securing and coordinating appropriate use of healthcare services, both by providing primary and preventive care and by acting as gatekeepers for secondary care services [1]. Previously, European GPs have reported high job satisfaction [2–7], largely attributed to high autonomy [8, 9] and good compatibility with family life [10]. However, a trend of increasing workload in general practice has been seen in several European countries [11, 12]. In England, studies report long and intense working hours, recruitment problems [13] and concerns for the well-being of both GPs and patients [14].

Several possible mechanisms explaining the increasing workload in general practice have been suggested [15]. In many European countries, healthcare reforms have transferred numerous tasks and responsibilities to primary care in order to reduce pressure on secondary care [16]. This implies that primary care now has increased responsibility for severely ill patients [17]. In addition, it has been suggested that new developments and treatment possibilities, as well as rising public expectations, have increased GPs' workload [15].

The Regular GP scheme was introduced in Norway in 2001. This list-based system entitles all inhabitants to register with a regular GP, and it has been regarded as one of the most successful public services in Norway [18], with high satisfaction among both patients and GPs [19–21]. Most GPs are self-employed, and the reimbursement system is based on a combination of capitation fees and fee-for-service [22]. About 10% of GPs are employed by their local municipality and get a fixed salary [23]. The regular GPs are responsible for coordinating healthcare services for the patients on their lists, and medical attestation and follow-ups for all absence from work of 3–8 days, including attestation for absence from high school. In 2012, a Coordination Reform was implemented, delegating more tasks to general practice. On average, a GP's patient list in Norway has approximately 1100 patients. This number has decreased in recent years [23], which may be a consequence of increased workload [24].

There is limited research on how increasing workload and the transfer of responsibilities to primary care may influence Norwegian general practice. This qualitative study aims to explore how GPs and their co-workers in Norway perceive and tackle their workload, and their experiences and reflections regarding explanations for and consequences of increased workload in general practice.

Material and method

Design

As this study is part of a project investigating different aspects of capacity pressure on health services [25], we wanted to identify possible mechanisms related to workload. We chose a qualitative method in order to explore and provide rich descriptions of these complex phenomena [26]. We applied a phenomenological approach, a methodology that relies on first-person accounts as the source of knowledge [27]. We collected data through interviews in urban and rural municipalities of Mid-Norway. We chose to conduct both focus groups and individual interviews for practical reasons, as not all of our participants in the same location could partake in interviews at the same time. In addition, we saw this as an opportunity to explore and compare dynamics when statements were given in groups as opposed to individual interviews.

Participants

The study participants were recruited by strategic sampling, via personal invitations by e-mail. We aimed to include GPs with varying sex, age, experience, size of practice, managing style and geographical location, thus securing a wide range of perspectives on the topic. To enlighten the topic further, we also included co-workers (health secretaries and nurses) from some of the practices in the interviews. A total of 23 GPs and 10 co-workers were interviewed, and the focus groups consisted of participants working at the same office. For the characteristics, see Table 1. Each participant cited is referred to with an individual number, as well as denoting the number of the focus group (G) in which they were interviewed or if they were interviewed individually (I).

Data collection

We conducted 11 interviews in Mid-Norway between September 2017 and January 2018 (Fig. 1). The interviews were held at the practices and lasted approximately 60 min. The authors alternated as the main interviewers, and at least one medical doctor participated in each interview. We used a semi-structured interview guide (Table 2), pilot tested by an academic GP. The interview guide was adjusted continuously throughout the study. Further, we followed up on statements made by previous participants, exploring if other participants shared the same experience. Interviews were audio recorded and transcribed verbatim by a secretary. All audio records were listened to and transcripts were anonymised, as well as being proofread by at least one of the authors. Interviews were reviewed throughout the study, and they continued until we agreed that sufficient information power was reached and no new themes were emerging [28].

Analyses

We used systematic text condensation, a thematic cross-case analysis based on Giorgi, developed and

Table 1 Characteristics of Participants

	N = 33
Sex	
Female	21
Male	12
Occupation	
GP	23
Co-worker	10
Age	
20–29	6
30–39	7
40–49	12
50–59	3
60–69	2
missing	3
Location	
Rural	16
Urban	17
GP characteristics	N = 23
Years as a GP	
< 2	5
2–4	1
5–9	7
10–19	8
≥20	2
List size	
< 900	3
900–999	4
1000–1099	5
1100–1199	3
1200–1299	4
1300–1399	1
1400–1499	1
≥1500	1
No list/intern	1
Speciality	
General practice[a]	13
Other	1
No	9
Days per week in the office	
2–3	6
4–5	17

[a]Completed 5 years of speciality training in general practice, and mandatory courses

modified by Malterud [29], to analyse the data. It consists of the following steps: 1) reading and listening to all the material and obtaining a total impression; 2) identifying, sorting and coding "meaning units", units of text providing knowledge of the phenomenon being studied; 3) condensing and abstracting the meaning within each of the codegroups; and, 4) synthesising the condensations into major topics and subtopics that reflect the interviewees' experiences of causes and consequences of their workload. The main research team consisted of one social scientist, one health economics scientist and four medical doctors, including one academic GP. They participated in all parts of the study, and read and coded the data material separately. Themes, content and coding were discussed thoroughly several times in a plenum, and adjusted by the research team. Academic GPs at our university were involved in the planning process of the study.

Ethics

No patient information was obtained in this study. The study was approved by the Regional Ethics Committee (2016/2158/REK Midt) and the Norwegian Data Inspectorate (54945). All participants signed a written consent to participate and were given the opportunity to withdraw from the study at any time.

Results

At the start of this study, our aim was to elucidate the participants' perceptions of their workload, and the potential explanations and consequences related to variations in workload. However, we noticed that the participants led the discussion into how their perceived workload had increased over the years. They further reflected on the mechanisms for this development. As this was a prominent feature throughout all of the interviews, we chose to let the participants elaborate on this, and integrated it in the further analyses of the material. We categorised the results into three main themes: (1) Heavy and increasing workload – more trend than fluctuation?; (2) Explanations for high workload; (3) Consequences of high workload.

Heavy and increasing workload – more trend than fluctuation?

Assuming fluctuations in workload, we asked the participants to identify what characterised periods of heavy workload. The participants all described variations in workload, both over weekdays and seasons. Both groups listed epidemics like influenza, with a higher inflow of patients, as resulting in increased workload. Particularly busy periods often occurred for GPs when they or their colleagues had a leave of absence or were preparing for or returning from one.

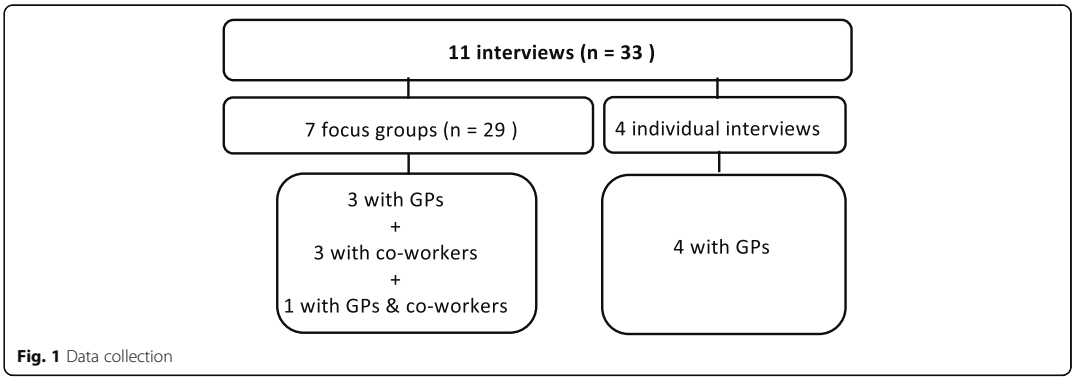


Fig. 1 Data collection

> *"It's almost like you can't be away for more than two days, because when you return, the pile of things to do almost feels impossible to handle."* I, female 1, GP

The co-workers, on the other hand, experienced higher workload when all the doctors at the office were present and thus there was a high turnover of patients. Furthermore, unplanned absence among the doctors was listed as a source of stress and increased workload for the co-workers, because they could not offer any appointments to the patients. The GPs reported now having longer working hours than before, and this despite many of them having reduced the number of patients on their lists. The participants were all experiencing heavy workload at the time of their interviews. GPs from both the focus groups and the individual interviews reported their current situation to be unsustainable.

> *"I think things can't go on like this. I have reached a threshold of what I can fulfil; I think something drastic has to change. […] You get so tired, because you're half an hour late all the time. It's like a 'rat race' really."* I, female 9, GP

However, GPs from two of the focus groups experienced their current workload as sustainable, despite increasing. They reflected upon this sustainability as being associated with the way they were organised. One of these practices was managed by the municipality, and the other had recently been reorganised, leaving the managerial position to a medical secretary. The GPs suggested that this allowed them more time for patient contact, as they were relieved from handling some of the administrative tasks such as financial matters, and sick leave among their co-workers.

Table 2 Original interview guide

- How would you describe your GP office?
  - o Compared to others?
- How is your working situation right now?
  - o How busy are you nowadays, in terms of workload?
- Describe a regular day, compared to a particularly busy day at work.
- Have you experienced situations that resulted in extreme time pressure? Which situations?
- What might be consequences of time pressure / increased workload for you?
- What kind of support do you get from your colleagues when you are out of time?
- Imagine a **day** when you were under particular time pressure
  - o How did you handle the situation?
  - o How did you prioritise?
- Imagine a **period** when you were under particular time pressure / had an increased workload.
  - o How did you handle the situation?
  - o How did you prioritise?
- Which patient groups take up most of your time or take a lot of time to treat?
- Considering your situation today, to what degree does time pressure / workload affect you?
- Do you have any suggestions on how to reduce the workload for GPs?
- What is your experience of cooperation between general practice and secondary care when patients are discharged from hospital?

Explanations for high workload

The participants reflected upon many possible explanations for the high and increasing workload. Notably, the contributory factors suggested as being most important varied among GPs within the same focus groups and when interviewed individually. The GPs often pointed to "local challenges", such as having many patients with complex issues, collaborating with the local hospital, and handling administrative and management duties. However, they hardly ever referred to how colleagues with similar challenges had handled these. The co-workers supported the GPs' explanations, but they also shared more general views on how they perceived societal developments as affecting their working conditions. Below, we give an overview of the mechanisms suggested as creating higher workload, divided into three prominent themes.

Transfer of tasks

The participants experienced increasing transfer of medical tasks from secondary to primary care. Follow-ups for patients with cancer and chronic conditions were generally perceived as meaningful, but also challenging and time-consuming. Many participants expressed vexation towards the transferral of more administrative tasks such as writing sick-leave certificates or transport requisitions related to their patients' hospital visits. The GPs experienced an increasing demand for new diagnostic investigations and tests, both prior to referral and after treatment in secondary care. They gave examples of discharge reports from secondary care instructing the GP to refer the patient to another specialist or radiological examination, thus causing extra workload. This was often perceived as a consequence of a more fragmented and subspecialised secondary care, focusing on shortening hospital stays, and it contributed to a feeling of impaired autonomy. Some GPs stated that they sometimes felt like they were working in "both primary and tertiary care", being expected to help patients with problems that could not be solved in secondary care.

> *"For instance, if we send a patient because of a stomach-ache, they do a gastroscopy, and if they don't find anything, they send him back, instead of taking care of the problem, like 'can it be something else?', and try to find out themselves, like they used to do before. Now they always bounce the ball back in our corner, and we have to do everything ourselves anyway!"* I, female 10, GP

Nevertheless, the GPs acknowledged that secondary-care professionals also have a high workload and do not necessarily intend to be condescending. Communication between primary and secondary care was commonly identified as challenging and time-consuming. This was a well-known problem, but was now perceived to have a higher impact on the workload, as the time pressure was higher. Difficulties in reaching and conferring with secondary-care professionals were thought to result in potentially unnecessary referrals.

> *"So I think many referrals could have been avoided, if they had time, and you didn't have to spend time in line on the phone."* G8, female 23, GP

Increased work per patient

The participants experienced an increasing amount of work per patient in recent years. Changes in legislation, developments in medicine, increasing investigation and treatment possibilities, a need for communication and cooperation with other parts of the healthcare system, and higher demand for documentation were all perceived as contributing factors.

> *"Something that has changed in very few years is that there is a lot more work to each patient. (…) Now there is a lot more we can do, (…), and then we had the Coordination Reform, with clearer commands in the discharge reports."* G8, female 21, GP

While some of the new tasks were regarded as important for patient care, others were perceived as meaningless and bureaucratic. An example frequently mentioned by GPs was writing health certificates.

> *"I don't need a medical degree to document that someone had a cough three days ago (...) nor to write a health certificate for parking needs for someone who has no legs (…) as doctors we have to do something reasonable."* G2, male 5, GP

The sum of these statutory tasks and demands was seen as a threat to the GPs' autonomy. The co-workers also reported that they were "writing and writing and writing" to document the work of their practice, although they did not believe this would improve patient health. There seemed to be a general consensus that administrative tasks and "paperwork" had increased considerably over the last decade:

> *"The workload comes mostly from the paperwork. I sit with paperwork until seven or eight o'clock every evening. I'm done with patients about four o'clock, so it's the paperwork that makes it impossible to pick up the kids, or cook dinner…"* I, female 1, GP

Changes in society

The participants reflected upon societal changes as explanations for the increasing workload. In general, both the GPs and co-workers experienced increasing patient expectations for healthcare services, treatment options, and their general health and well-being. The co-workers suggested that a lack of family support and limited social networks often resulted in an increased number of doctor visits. They gave examples of minor issues that previously could be solved by "asking grandmother".

> *"People see their GP much more often nowadays. (…) Now you see the doctor at once – if you've been feeling ill for a few days (…) If a child gets a rash, then the parents go straight to the doctor to check it out. They*

didn't do that before. Now they demand an answer – 'What is this?'" I, female 10, GP

Some of the younger GPs suggested that the feeling of time pressure throughout the day resulted in many GPs preferring not to work as many hours as they had previously, similar to others in the society. On the other hand, some of the more experienced co-workers thought the doctors worked even more now and had higher competence in meeting patients' expectations.

"Today's GPs are different to those of 20 years ago. Before, they were mostly elderly, and men. Now, there are many women, and many young people with kids and completely different priorities. They want to go home at a decent time, pick up the kids, make dinner and drive to football practice." G2, male 2, GP

The GPs reported that they experienced administrative and economic duties in the GPs' offices to have become more advanced and complicated in the latest years. They perceived it as more demanding to handle employer responsibilities, such as dealing with pensions and sick leave for their staff. In addition, the expenses for running their offices had increased in recent years due to, e.g., increased requirements for electronic equipment and salaries for employees. As one experienced GP said:

"It has changed totally. And the capitation fee covers less and less of our real expenses. (…) I used to do my own accounting, but now I can't possibly do it, because so much has changed. It's more like running a company. That's not what I intended to do (laughs). So considering this, it was much easier to be publicly employed." I, female 10, GP

This caused economic worries for the GPs, and prevented them from reducing their patient lists and, hence, their workload, because parts of their financing are based on the size of their patient lists.

Consequences of high workload

Both the GPs and co-workers expressed that they now perceived busy days as the "new normal". As a response to this, the GPs said they were forced to adjust their way of working by prioritising harder. They prioritised patient consultations, postponed documentation and administrative work to evenings and weekends, and were left with little time for personal rest and recuperation. System-level work, such as participating in meetings, forums and other arrangements at the municipal level, and preventive care were given less priority due to lack of time. Further, the GPs expressed worries about their professional development being negatively impacted through, for instance, postponing or skipping educational courses.

Consequences for patients and the healthcare system

Both GPs and co-workers described how high workload had general consequences for patients, such as longer waiting times for appointments, reduced continuity of care due to use of locums, and possibly reduced patient satisfaction. They also shared their thoughts regarding how high workload could lead to suboptimal handling of some patient groups, such as patients with chronic illnesses or complex problems, the elderly, patients with mental health problems and patients with a minority background.

"It does affect the patients, definitely – regarding waiting times, availability, phone calls and, to some extent, the treatment and the care they receive." G2, male 3, GP

We found three different perspectives among the GPs regarding whether and how heavy workload influenced their own clinical decision-making, such as diagnostics, referrals, prescriptions and sick leave. All three perspectives were generally represented by different individuals within the focus groups. There did not seem to be any consistency regarding how these perspectives were related to GPs' characteristics such as age, experience, gender, or geographical location.

The first perspective was that heavy workload definitely influenced clinical decisions. The GPs gave examples of a lower threshold for prescribing antibiotics to children, and for referring patients with conditions that could have been treated in general practice, such as excessive ear wax and potential deep vein thrombosis. As a female GP said:

"We try not to do it. We are all quite experienced here, but it's hard to resist when there is so much to do." G7, female 14, GP

The GPs reflected upon how this could have paradoxical effects, and in turn cause more work for the healthcare system and themselves. They were conscious of their gatekeeper function, and described increased referrals as an unfortunate trend they wished to avoid.

"The more time you have, and the better you're feeling, both in private and at work, the more guts you have to keep calm and unaffected, which is the art of general practice. And then it's the gatekeeper function – we have to make sure we don't refer too many patients – both for the sake of the patients and for the community." G8, female 21, GP

The second perspective we found was that heavy workload partially influenced clinical decisions. These GPs were worried that time pressure affected how they interacted with the patients and increased their tendency to take resource-demanding shortcuts in medical investigations. They proposed that stress throughout the work day could increase the risk of making mistakes, or prioritising incorrectly. However, they did not believe that decisions such as referring patients were affected.

> *"It's about being present in the consultation. When in a hurry, you keep more distanced. Maybe you try to find some shortcuts to get things done in a shorter time."* G10, female 31, GP

The third perspective among the GPs was that heavy workload did not influence clinical decision-making at all, and that the patients were not affected directly.

> *"We have to state that the patient is our first priority, and that's why our days look like they do."* G2, male 4, GP

All GPs expressed their belief that the trend of heavy and increasing workload had negative impacts on the healthcare system, especially through recruitment problems in general practice.

> *"I think it is a symptom that we can't recruit enough GPs, and then there will be a huge problem in some years." G2, male 5, GP*

Personal consequences for GPs

GPs with children expressed problems regarding combining their job with family life. Many experienced difficulties in getting to kindergarten or school before closing time, finding time to eat dinner with their family, and taking part in their children's recreational activities. These GPs underlined the importance of having a partner with flexible working hours, so that they could stay at the office for as long as required. GPs without family responsibilities said they felt this was an advantage when the workload was high, and that they could relieve their colleagues when needed.

The GPs described that the workload had consequences for their own health and well-being. At work, they often skipped coffee breaks, shortened their lunch break, and postponed toilet visits. At home, some said that they felt exhausted, easily irritated and stressed, and did not find time to exercise. Two of the younger female GPs worried about being burned out, and not being able to continue working as a GP in the future.

> *"Maybe I can stay another year or so, because I love my job. […] I just need some space to breathe in my working day; otherwise I think I will burn out."* I, female 9, GP

Similarly, the co-workers also felt stressed at work when the workload was high. However, in contrast to the GPs, they also highlighted how they did not have to bring this stress home with them, and they spoke positively of their regulated work hours. The GPs described their decreased motivation for continuing with their job, and a young male doctor said that, based on his experiences of the last year, he no longer wanted to be a GP. Several GPs had considered quitting or were looking for other jobs. Nevertheless, all of our participating GPs expressed a genuine love for their work, felt that their job was meaningful, and wished that conditions would improve so that they could continue.

> *"Yes, it's a wonderful job where you meet all these incredibly nice people that you wouldn't have met otherwise. It is varied, gives lots of challenges, both in medical and organisational terms. (…) In many ways, it is the best job in the world – you even have an illusion of autonomy (laughs)."* G10, male 33, GP

Discussion

Key findings

Our main finding was that the participants perceived the workload in general practice as heavy and having substantially increased in recent years. The suggested explanations for and consequences of heavy workload seemed to vary among the GPs, but they all experienced an increased workload per patient. The participants expressed concerns for the future in regards to patient safety, GPs' health and motivation, and the recruitment of new GPs.

Strengths and limitations

A strength of this qualitative approach was that our research group consisted of researchers both with and without clinical experience from general practice. We believe that this balanced our preconceptions, subjective views, and experiences related to clinical practice. It may have also enabled us to recognise various aspects of the topic and potentially led to a more thorough understanding of both the research question and the material. There was little difference between the opinions expressed in focus groups or in individual interviews. Including the GPs' co-workers, both in separate focus groups and in a focus group together with GPs, gave us a nuanced view of the topic.

Qualitative studies have known limitations concerning transferability. We included only GPs' offices in

Mid-Norway, as inclusion of participants from a larger geographical area was not feasible within the scope of this study. Among all of the GPs asked to participate, only four declined our request, as they could not find time for the interview.

In the months between planning the study and conducting the interviews, there was substantial media attention regarding workload in general practice in Norway. This might have influenced the way we asked questions during the interviews, as well as the way we interpreted the material. It might also have affected the respondents' views and thoughts about their working conditions and the workload in their practice, and may possibly have led to a polarisation of the opinions.

Comparison with existing literature

A clear finding was a perception of heavy and increasing workload. Some of the GPs in our study suggested that this perception could partly be influenced by changes in their mentality and expectations. While being a doctor was previously considered a lifestyle choice, today's young doctors often see it 'merely' as a job [30].A study among doctors working at Norwegian hospitals found this difference in perspectives to be associated with an increased work–life imbalance [31]. Several of the GPs in our study reported a work–life imbalance, particularly those with family responsibilities.

However, our findings of GPs experiencing high and increasing workload are supported by recent statistics. Norwegian GPs' weekly working hours increased by 7 h from 2014 to 2018, resulting in an average of 56 h. Approximately 50% of GPs report that they are working during weekends, even when off duty [32]. At the national level, the total number of patient consultations in general practice is rising [33]. Simultaneously, the number of patients on GPs' lists has decreased [23], which can be interpreted as more work per patient.

Similar trends of increasing workload have also been reported in other European countries [34]. In England, several studies have reported increased workload for GPs [14, 15, 35], although one report found a slight reduction in working hours between 2012 and 2015 [36]. In Denmark, both workload and working hours in general practice have increased substantially [37]. In a survey of 25 EU countries, 19 reported that "workload in general practice is unreasonable and unsustainable" [38]. The countries that reported general practice workload as reasonable had a common feature of GPs working 8 h or less per day. In comparison, only 10% of the GPs in Norway have weekly working hours within the Norwegian "norm" of 37.5 [32]. Noticeably, the co-workers in our study spoke positively of their regulated working hours as a counterbalance to the increased stress at work, and there is no known recruitment problem in this profession.

In sum, we found a wide range of possible explanations for the increasing and high workload. This is an important finding, as it implicates the complexity of feasible approaches to relieve the situation. Many of the explanations presented by our participants regarding patient and system factors are described in previous research and reports from other countries [15], and could be seen as part of a more general societal development. Many of the mechanisms are also difficult, if at all possible, to change. One cannot stop the population from ageing, or remove multimorbidity. Changing the public's expectations of what the health care system can and should help them with is probably needed [39], but this will take time. Although some medical technology is held back for economic and ethical reasons, the trend of more diagnostics and treatment possibilities is not easy to halt or reverse [40]. In our study, the participants suggested to reduce the number of patients per regular GP in order to reduce the workload. We believe this could relieve the workload, but requires recruitment of a large number of new GPs, which constitutes further public expenditure. On the other hand, if the Regular GP scheme is weakened, this could also potentially result in higher expenditure, as a well-functioning primary health care in general is shown to be crucial for public health at a lower cost for the society [41]. Further, continuity of care have been associated with both lower mortality [42], and lower use of secondary health care services [43].

In response to the increasing workload, the GPs in our study handled the situation by for instance expanding working hours and increasing the number of GPs in their office. They pointed out that most of these changes were only temporarily useful. However, outsourcing the position as daily manager had relieved two of the offices from increasingly administrative and employer duties. We were surprised that the participants seemed to have little knowledge about the different possibilities in for example management forms. We believe sharing experiences like these could be helpful for other GPs.

We found that the GPs had different views on if and to what extent patients and the health care system were affected by heavy workload. It was prominent how they stated that they were willing to go to great lengths to prevent their patients from being directly affected by the heavy workload, and explained this as an important reason for the trend of longer working hours. Previous research has shown that GPs are able to adapt to higher workload during periods of higher demand, like influenza pandemic [44]. However, our participants strongly pointed out that they perceived the current situation

with high workload to be a trend more than periodic variation. Exposure to high workload over time, increases the risk of burnout [45, 46], which in turn is shown to be harmful for patient safety [47, 48], and also suggested to be associated with higher referral rates [49]. To our knowledge, there are no recent studies describing the prevalence of burnout among GPs in Norway. Nevertheless, European studies report an increasing number of doctors being burned out [36, 50], and suggests association with increased workload.

Autonomy has earlier been reported as a motivation for choosing a career in general practice [10]. If our finding of impaired autonomy among the GPs is a widespread phenomenon, this may impact the recruitment of GPs. Excessive working hours in general practice have also been suggested to cause lower job satisfaction and give recruitment problems [13]. Taking into account that more than one-third of the GPs in Norway are over the age of 55 [23] along with the recruitment challenges in both rural and urban areas [51], the Norwegian healthcare system and wider society face potential challenges.

Despite their heavy workload, the participants were still enthusiastic about their work and societal responsibilities. Nevertheless, they all had concerns for their future, the Regular GP scheme and, like their colleagues in England [14], the recruitment of new GPs. They perceived the Regular GP scheme's current situation as unsustainable, and expressed worries that they will not be able to provide the population with the expected level of service for primary care in the future [52], despite the Regular GP scheme previously being regarded as a great success [19–21].

Conclusions

This study found heavy and increasing workload in general practice in Norway. The explanations appear to be multi-faceted and many are difficult to reverse. The GPs expressed worries that they will not be able to provide the population with the expected care and services in the future.

Abbreviations

GP: General practitioner

Acknowledgements

The authors thank all participants who gave their time to take part in the study, academic GPs at the Department of Public Health and Nursing for useful inputs in the planning phase of the study, and professors Linn O. Getz and John-Arne Skolbekken for valuable help in the writing process.

Funding

This work is part of the research project "Pressure in the health care systems, consequences for efficiency, patient flow and patient safety" funded by the Norwegian Research Council (grant number 256579).
It was also supported by the Department of Public Health and Nursing at the Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology.

Availability of data and materials

The datasets generated and analyzed during the current study consists of transcribed material from focus groups and individual interviews, and are not publicly available due to the consideration of the privacy and consents of the study participants, but are available from the corresponding author on reasonable request.

Authors' contributions

All authors have contributed in the planning of the study, literature search and writing the manuscript. The interviews were conducted by authors ERS, LJS, KP, MTM, and SLK. Analysis were performed by all authors except JHB. All authors read and approved the final manuscript.

Ethics approval and consent to participate

No patient information was obtained in this study. The study was approved by the Regional Ethics Committee (2016/2158/REK-midt) and the Norwegian Data Inspectorate. All participants signed a written consent to participate, and were given the opportunity to withdraw from the study at any point.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Author details

[1]Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Håkon Jarls gate 11, 7030 Trondheim, Norway. [2]SINTEF Digital, Department of Health, Trondheim, Norway. [3]Forensic Department and Research Centre Bröset, St. Olav's University Hospital Trondheim, Trondheim, Norway.

Received: 7 December 2018 Accepted: 26 April 2019
Published online: 21 May 2019

References

1. Starfield B. Is primary care essential? Lancet. 1994;344:1129–33.
2. Groenewegen PP, Hutten JB. Workload and job satisfaction among general practitioners: a review of the literature. Soc Sci Med. 1991;32:1111–9.
3. Hetlevik O, Hunskar S. The length of the patient list, waiting lists, workload and job satisfaction among general practitioners in Bergen. Tidsskr Nor Laegeforen. 2004;124:813–5.
4. Iversen L, Farmer JC, Hannaford PC. Workload pressures in rural general practice: a qualitative investigation. Scand J Prim Health Care. 2002;20:139–44.
5. Moth G, Vestergaard M, Vedsted P. Chronic care management in Danish general practice--a cross-sectional study of workload and multimorbidity. BMC Fam Pract. 2012;13:52.
6. Brondt A, Vedsted P, Olesen F. General practitioners' job satisfaction. Ugeskr Laeger. 2007;169:2521–5.
7. Gotz K, Broge B, Willms S, Joos S, Szecsenyi J. Job satisfaction of general practitioners. Med Klin (Munich). 2010;105:767–71.
8. Nylenna M, Aasland OG. Job satisfaction among Norwegian doctors. Tidsskr Nor Laegeforen. 2010;130:1028–31.
9. Aasland OG, Rosta J, Nylenna M. Healthcare reforms and job satisfaction among doctors in Norway. Scand J Public Health. 2010;38:253–8.
10. Roos M, Watson J, Wensing M, Peters-Klimm F. Motivation for career choice and job satisfaction of GP trainees and newly qualified GPs across Europe: a seven countries cross-sectional survey. Educ Prim Care. 2014;25:202–10.
11. Dayan M, Arora S, Rosen R, Curry N. Is general practice in crisis? London: Nuffield Trust; 2014. https://www.nuffieldtrust.org.uk/resource/is-general-practice-in-crisis?gclid=EAIaIQobChMI1_nR1YmU3AIVQuAZCh1NoQ2CEAAYASAAEgL-qPD_BwE. Accessed 10 Oct 2018
12. Pedersen AF, Nørøxe KB, Bro F, Vedsted P. Alment praktiserende lægers psykiske arbejdsmiljø og jobtilfredshed 2016. Aarhus: Forskningsenheden for Almen Praksis, Aarhus Universitet; 2016. http://feap.au.dk/fileadmin/feap/Rapport_Praktiserende_laegers_arbejdsmiljoe_30.10.2016.pdf. Accessed 10 Oct 2018

13. Marchand C, Peckham S. Addressing the crisis of GP recruitment and retention: a systematic review. Br J Gen Pract. 2017;67:e227–e37.
14. Croxson CH, Ashdown HF, Hobbs FR. GPs' perceptions of workload in England: a qualitative interview study. Br J Gen Pract. 2017;67:e138–e47.
15. Baird B, Charles A, Honeyman M, Maguire D, Das P. Understanding pressures in general practice. London: The King's fund; 2016.
16. Grimsmo Anders MJ. Norwegian coordination reform in an international perspective [Norsk samhandlingsreform i et internasjonalt perspektiv]. In: Institute of public health and nursing. The Research Counsil of Norway: Norwegian University of Science and Technology; 2015. https://app.cristin.no/results/show.jsf?id=1262523.
17. omsorgsdepartementet] TmoHaCsH-o. The Coordination Reform. Proper treatment - at the right place and right time [Samhandlingsreformen. Rett behandling - på rett sted - til rett tid. St.meld.nr. 47 (2008–2009)]. In. Oslo; 2009.
18. Åmotsbakken MD. The regular GPs gain popularity in the population [Fastlegene stadig mer populære i befolkningen]. Oslo: The Norwegian Medical Association [Den norske legeforening]; 2017. http://legeforeningen.no/Nyheter/2017/Fastlegene-stadig-mer-populare-i-befolkningen/. Accessed 3 July 2018
19. Grytten J, Skau I, Carlsen F. Patient satisfaction with general practice services before and after the list patient reform. Tidsskr Nor Laegeforen. 2004;124:652–4.
20. Grytten J, Skau I, Sorensen R, Aasland OG. Physicians' work situation a year after the list patient system reformTidsskr Nor Laegeforen. 2004;124:358–61.
21. Sorensen M. What do we know about patient satisfaction with the Regular GP scheme? The Tidsskr Nor Laegeforen. 2018;138(10). https://doi.org/10.4045/tidsskr.17.0861.
22. Eide TB, Straand J, Bjorkelund C, Kosunen E, Thorgeirsson O, Vedsted P, et al. Differences in medical services in Nordic general practice: a comparative survey from the QUALICOPC study. Scand J Prim Health Care. 2017;35:153–61.
23. The Norwegian Directorate of Health H. Regular GP statistics (Fastlegestatistikk). In. https://www.helsedirektoratet.no/statistikk/statistikk/fastlegestatistikk#fastlegestatistikk-2017 2017.
24. Norway S. Shorter patient lists, longer working days? [Kortere pasient lister, lengre arbeidsdager?]; 2018.
25. Nilsen S. Health care services under pressure - consequences for patient flows, efficiency and patient safety in Norway. 2019. https://www.researchgate.net/project/Health-care-services-under-pressure-Consequences-for-patient-flows-efficiency-and-patient-safety-in-Norway. Accessed 3 Mar 2019.
26. Sofaer S. Qualitative methods: what are they and why use them? Health Serv Res. 1999;34:1101–18.
27. Mishler EG. Research interviewing: context and narrative. Cambridge: Harvard University Press; 1986.
28. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res. 2016;26(13):1753–1760. https://doi.org/10.1177/1049732315617444.
29. Malterud K. Systematic text condensation: a strategy for qualitative analysis. Scand J Public Health. 2012;40:795–805.
30. Hertzberg TK, Skirbekk H, Tyssen R, Aasland OG, Ro KI. The good doctor - strong and persevering. Tidsskr Nor Laegeforen. 2016;136:1631–4.
31. Hertzberg TK, Skirbekk H, Tyssen R, Aasland OG, Ro KI. The hospital doctor of today - still continuously on duty. Tidsskr Nor Laegeforen. 2016;136:1635–8.
32. Rebnord IK, Eikeland OJ, Hunskår S, Morken T. Regular GPs' time use. Bergen: Nasjonalt kompetansesenter for legevaktmedisin, Uni Research Helse; 2018. https://www.regjeringen.no/no/dokumenter/fastlegers-tidsbruk/id2592992/. Accessed 3. July 2018
33. Norway S. Increased use of GPs in 2017 [Vekst i bruk av fastlegene i 2017]. 2017. https://www.ssb.no/helse/artikler-og-publikasjoner/vekst-i-bruk-av-fastlegene-i-2017. Accessed 29 Nov 2018.
34. Gyorffy Z, Dweik D, Girasek E. Workload, mental health and burnout indicators among female physicians. Hum Resour Health. 2016;14:12.
35. Dale J, Potter R, Owen K, Leach J. The general practitioner workforce crisis in England: a qualitative study of how appraisal and revalidation are contributing to intentions to leave practice. BMC Fam Pract. 2016;17:84.
36. PRU Comm GJ. Eighth national GP worklife survey report; 2015.
37. Pedersen AFNK, Bro F, Vedsted P. GPs' psychological work environment and job satisfaction 2016 [Alment praktiserende lægers psykiske arbejdsmiljø og jobtilfredshed 2016]. In: Forskningsenheden for Almen Praksis, Aarhus Universitet 2016; 2016.
38. McCarthy M. Sustainable general practice: looking across Europe. Br J Gen Pract. 2016;66:36.
39. Thompson M, Walter F. Increases in general practice workload in England. Lancet. 2016;387:2270–2.
40. Hofmann BM. Too much technology. BMJ. 2015;350:h705.
41. Starfield B, Shi L, Macinko J. Contribution of primary care to health systems and health. Milbank Q. 2005;83:457–502.
42. Pereira Gray DJ, Sidaway-Lee K, White E, Thorne A, Evans PH. Continuity of care with doctors-a matter of life and death? A systematic review of continuity of care and mortality. BMJ Open. 2018;8:e021161.
43. Hansen AH, Halvorsen PA, Aaraas IJ, Forde OH. Continuity of GP care is related to reduced specialist healthcare use: a cross-sectional survey. Br J Gen Pract. 2013;63:482–9.
44. Simonsen KA, Hunskaar S, Sandvik H, Rortveit G. Capacity and adaptations of general practice during an influenza pandemic. PLoS One. 2013;8:e69408.
45. Linzer M, Visser MR, Oort FJ, Smets EM, McMurray JE, de Haes HC. Predicting and preventing physician burnout: results from the United States and the Netherlands. Am J Med. 2001;111:170–5.
46. Maslach C, Leiter MP. Chapter 43 - burnout. In: Fink G, editor. Stress: concepts, cognition, emotion, and behavior. Edn. San Diego: Academic Press; 2016. p. 351–7.
47. Hall LH, Johnson J, Watt I, Tsipa A, O'Connor DB. Healthcare staff wellbeing, burnout, and patient safety: a systematic review. PLoS One. 2016;11:e0159015.
48. Panagioti M, Geraghty K, Johnson J, Zhou A, Panagopoulou E, Chew-Graham C, et al. Association between physician burnout and patient safety, professionalism, and patient satisfaction: a systematic review and meta-analysis. JAMA Intern Med. 2018;178:1317–30.
49. Kushnir T, Greenberg D, Madjar N, Hadari I, Yermiahu Y, Bachner YG. Is burnout associated with referral rates among primary care physicians in community clinics? Fam Pract. 2014;31:44–50.
50. Soler JK, Yaman H, Esteva M, Dobbs F, Asenova RS, Katic M, et al. Burnout in European family doctors: the EGPRN study. Fam Pract. 2008;25:245–65.
51. Sæther AS. [Lack of regular GPs in Tromsø and Bergen] Fastlegemangel i Tromsø og Bergen. Oslo: Verdens Gang. 2018. https://www.vg.no/nyheter/innenriks/i/e1E1pR/fastlegemangel-itromsoe-og-bergen. Accessed 4 July 2018.
52. Wonka Europe. The European definition of general practice/Family Medicine. 2011. http://www.woncaeurope.org. http://www.woncaeurope.org/sites/default/files/documents/Definition%203rd%20ed%202011%20with%20revised%20wonca%20tree.pdf. Accessed 3 Mar 2019.

Paper II

Family Practice, 2021, 1–8
https://doi.org/10.1093/fampra/cmab120

OXFORD

Health Service Research

Effects of GP characteristics on unplanned hospital admissions and patient safety. A 9-year follow-up of all Norwegian out-of-hours contacts

Ellen Rabben Svedahl[1,*], Kristine Pape[1], Bjarne Austad[1], Gunnhild Åberge Vie[1], Kjartan Sarheim Anthun[1,2], Fredrik Carlsen[3], Johan Håkon Bjørngaard[1,4]

[1]Department of Public Health and Nursing, Norwegian University of Science and Technology, Trondheim, Norway, [2]Department of Health Research, SINTEF Digital, Trondheim, Norway, [3]Department of Economics, Norwegian University of Science and Technology, Trondheim, Norway, [4]Faculty of Nursing and Health Sciences, Nord University, Levanger, Norway

*Corresponding author: Department of Public Health and Nursing, Norwegian University of Science and Technology, PO Box 8905, 7491 Trondheim, Norway. Email: ellen.r.svedahl@ntnu.no

Abstract

Background: There are substantial differences in hospital referrals between general practitioners (GPs); however, there is little research on the consequences for patient safety and further healthcare use.

Objective: To investigate associations between out-of-hours GP characteristics, unplanned hospital admissions, and patient safety.

Methods: This cohort study included all Norwegian out-of-hours services contacts from 2008 to 2016, linked to registry data on patient characteristics, healthcare use and death, and GP age, sex, specialist status, out-of-hours service experience, and prior admission proportion. We estimated the impact from GP characteristics on (i) immediate unplanned hospital admissions for "all conditions," (ii) immediate unplanned hospital admissions for "critical conditions," (iii) 30-day unplanned hospital admissions, (iv) 30-day hospital costs, and (v) 30-day risk of death. To limit confounding, we matched patients in groups by age, time, and location, with an assumption of random assignment of GPs to patients with this design.

Results: Patients under the care of older and male GPs had fewer immediate unplanned hospital admissions, but the effects on cumulative 30-day unplanned hospital admissions and costs were small. The GPs' prior admission proportion was strongly associated with both immediate and 30-day unplanned hospital admissions. Higher prior admission proportion was also associated with admitting more patients with critical conditions. There was little evidence of any associations between GP characteristics and 30-day risk of death.

Conclusions: GPs' prior admission proportion was strongly associated with unplanned hospital admissions. We found little effects on 30-day mortality, but more restrictive referral practices may threaten patient safety through missing out on critical cases.

Key Messages

- GP characteristics were associated with substantial differences in referrals.
- Older, male, and specialist GPs were associated with lower odds for referrals.
- The GP's referral history was the most important determinant for further referrals.
- Reducing referrals may threaten patient safety through missing critical cases.
- The differences in referrals had minor effects on 30-day mortality.

Lay Summary

Referral for specialized health services is a key part of the general practitioner (GP) role. Differences in referrals between primary care physicians have been widely studied, as they represent a target for reducing the use of specialized health services. However, the potential consequences beyond the actual referral have received little attention. Studying associations between physician characteristics and clinical decisions are difficult because physicians often systematically see different patient populations with different morbidity. Previous findings showing large differences in clinical decisions regarding referrals and hospital admissions may suffer from confounding. With our carefully matched study design, we could assume that the assignment of physicians to patients was random. We found substantial differences in referrals associated with GP characteristics. Seeing older and male GPs and specialists in family medicine were associated with fewer immediate unplanned hospital admissions but did not substantially influence unplanned hospital costs within 30 days. However, GPs with a history of admitting many of their recent patients had a substantial higher tendency to admit their future patients and represented a higher use of health services and costs. These GPs also referred more critically ill patients, an essential aspect of patient safety. The differences in referrals had minor impact on the patients' 30-day risk of death.

Key words: after-hours care, general practitioners, health services research, patient admission, patient safety, referral and consultation

Background

In most healthcare systems, the overall general policy is to reduce unnecessary and unplanned hospital admissions as they are demanding for the health services, costly for society, and may increase the risk of overtreatment and complications for the patients. Gatekeeping in primary care is shown effective to control the use of specialized health services.[1] Thus, evidence suggesting large differences in referral practices between primary care physicians highlight a target for quality improvement and reduction of unnecessary admissions.[2–7] However, this evidence may suffer from potential unmeasured confounding from different patient populations between the physicians because important patient characteristics may not be readily available in such studies.[8] Hence, what appears as differences in physicians' referral patterns may instead reflect differences in their patients' healthcare needs. In health systems where regular general practitioners (GPs) mainly serve a selected patient population, strong associations between GP characteristics and patient characteristics may be observed, without this reflecting real differences in referral threshold. However, the out-of-hours setting, where the patients to a lesser degree choose their GP, may be better suited for studying such associations. Out-of-hours medical services provide urgent primary medical care outside office hours and hold an essential gatekeeping role for unplanned hospital admissions.[9,10] GPs are the backbone in the out-of-hours services in many countries.[11] Increasing pressure on primary care is now challenging the contribution from experienced GPs and may also lead to a shift in the characteristics in the GPs staffing both normal hours and out-of-hours primary care.[12,13] Knowledge about the potential effects of GP characteristics on referral differences is therefore valuable. Further, there is a lack of research on the consequences of such differences for patient safety and healthcare use.

The aim of this study was to investigate the impact of GP characteristics on unplanned hospital admissions. To handle confounding from different patient populations, we studied patient contacts in the out-of-hours setting, and further matched patients in comparable groups. In addition to looking at the differences in immediate unplanned hospital admissions for all conditions, we included outcomes reflecting patient safety; immediate admissions for critical conditions, 30-day hospital use and costs, and 30-day risk of death.

Materials and methods

Study setting

The Norwegian out-of-hours services is a statutory municipal service, organized as a GP cooperative which is the most dominant model in Europe.[11] Other physicians also staff the out-of-hours services, but GPs contribute with about half of the contacts.[14] Most acute illness outside office hours are handled in the out-of-hours service as a primary care emergency unit, and patients are assigned to the physician on-call in their area. See Supplementary Material for details of the study setting.

Study population

This study is based on complete national data on patient contacts with primary care physicians from the Control and Payment of Health Reimbursement Register (KUHR).[15] The study population comprises all patients contacting the Norwegian out-of-hours services in the period 2008–2016, assessed by physicians also

working as regular GPs during office hours. We included out-of-hours services contacts between 16:00 and 08:00 on weekdays and whole Saturday, Sunday, and public holidays. We used a unique identification number to link patient data to somatic hospital visits in the Norwegian Patient Registry,[16] demographical information including municipality code, immigration, and education status from Statistics Norway,[17] and date of death from the Norwegian Cause of Death Registry.[18] Unplanned admissions to psychiatric care were not included in this study. GP characteristics available from the Norwegian General Practitioner Register were linked to each patient contact by a unique physician ID.[19] See Supplementary Material for details of the study population and data sources.

Study design

We designed our study so we could arguably assume that the patients' measured and unmeasured pretreatment conditions, were balanced between the compared patient groups and independent of the characteristics of the GP(s) on-call. This implied a matching procedure, defined by combining information on patients: (i) being in same 10-year age groups, (ii) visiting the same out-of-hours service, (iii) on the same weekday in the same month and year, and (iv) in the same 8-h time unit during the day (Fig. 1). By matching patients and analyzing only within-group variability, we effectively controlled for all confounding that was constant within each patient-matched group. As an example, we compared patients in a given 10-year age group visiting a particular out-of-hours service on a Tuesday evening in January 2015 with patients in the same 10-year age group visiting the same service one of the three other Tuesday evenings in January 2015. For about 70% of patients, only one GP was available on-call (in the current 8-h time unit and age group). To avoid the effect of possible patient selection in situations where two or more GPs were on-call at the same time, we used the weighted average of GP characteristics within each 8-h time unit in each service.

We excluded out-of-hours claims where the patient's regular GP was present at the out-of-hours service, as these contacts could easily be made based on an agreement between the patient and the GP. See Supplementary Material for details of the study design.

Outcome variables

The study had the following outcomes:

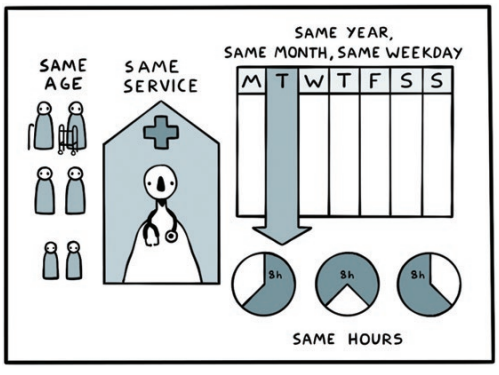


Fig. 1. Study design. Comparable groups were made by matching patients in the same 10-year age group, visiting the same out-of-hours service, the same weekdays in the same month and year, and the same 8-h time unit.

1) Immediate unplanned hospital admissions, defined as urgent hospital contact registered within 10 h:

   a. Admissions for all conditions vs not admitted.
   b. Admissions for critical conditions vs not. Critical conditions were measured as contacts resulting in a severe discharge diagnosis, such as myocardial infarction, stroke, pulmonary embolism, severe head injuries, fractures, and infections (see Supplementary Material for a complete list of ICD-10 codes).

2) Thirty-day unplanned hospital use presented per 1,000 GP contacts:

   a. Cumulative incidence of unplanned hospital admissions.
   b. Cumulative costs from unplanned hospital stays starting within 30 days after the index contact. The costs were calculated from diagnosis-related group points.[20]

3) Thirty-day risk of death (only for the two oldest patient groups due to few deaths among the youngest).

Exposure variables

GP characteristics at the time of each contact included the GPs' sex, age, and speciality status. Further, we measured the GPs' previous working experience from out-of-hours services (defined as "low" if less than 200 out-of-hours contacts during the two preceding years) and the "prior admission proportion," calculated as the proportion of out-of-hours contacts during the preceding 4-month period resulting in immediate unplanned hospital admission, excluding the contact with the index patient. We divided the study population into four equal sized groups (quarters), based on the prior admission proportion of the GP(s) on-call. The top quarter was patients under the care of GPs with the highest hospital referral tendency, and the bottom quarter was under the care of GPs with the lowest. By comparing the top with the bottom quarters, we avoided comparisons with extreme deviations from normal practice.

Statistical analyses

Patient contacts were matched in groups as described above. All multivariable analyses were performed with a within-matched-group estimator. We performed separate analyses and present results for three age groups: 0–10 years, 11–69 years, and 70 years and older. Immediate unplanned hospital admissions and 30-day risk of death were analyzed with a within-matched-group estimator with conditional logistic regression (clogit command in Stata), while the 30-day unplanned hospital admissions and costs (for hospital stays starting 0–30 days after a contact) were estimated using a within-matched-group estimator with linear regression (xtreg, fe in Stata). In addition to the matching procedure, all estimates were adjusted for patient sex, age, and age squared. Precision was evaluated with robust 95% confidence intervals (CIs). The analyses were performed with Stata version 15.1.

Assumptions and additional analyses

Within each matched group, we assumed that the GP characteristics would not be associated with possible confounding characteristics of the patients. To justify this assumption, we performed balance tests calculating the associations between potentially confounding patient characteristics and the GP characteristics. The patient characteristics included age, sex, immigration status (yes/no), and education

(completed less than 13 years). Further, as a proxy for patient morbidity, we used the patients' health service contacts 30 days before each contact (i.e. GP visits, planned and unplanned hospital admissions, and outpatient visits), in addition to a Charlson Comorbidity Index score based on diagnoses from the most recent hospital visit the previous month.[21] Results are presented in Supplementary Table 1. In Supplementary Table 4, results for the main analysis are presented for each 10-year age group. We also analyzed all exposure variables adjusted for each other (Supplementary Fig. 1). Further, we compared the OR for immediate unplanned hospital admission for patients under the care of the active GPs and the other physicians (defined as non-GPs) staffing the out-of-hours services (Supplementary Table 3).

Results

We present descriptive results in Table 1. In the age group 0–10 years, 6.2% had an unplanned admission to hospital within 10 h of the index contact. For the age group 11–69 years and 70 and older, the corresponding numbers were 12.4% and 25.8%. For patients aged 11–69 years, 0.2% died within 30 days after the index contact. For patients aged 70 years and older, 4.6% died.

Immediate unplanned hospital admissions

All conditions

A 10-year increase in the GP age was associated with 5%–8% reduced odds of unplanned hospital admission (adjusted odds ratio [aOR] 0.92, 95% CI 0.90–0.93 in patients 0–10 years, aOR 0.95, 95% CI 0.94–0.96 in patients 11–69 years, and aOR 0.94, 95% CI 0.93–0.95 in patients 70 years and older) (Fig. 2). Contact with a male vs female GP gave 12% lower odds for hospital admission for patients aged 70 years and older and 24% lower odds for patients aged 0–10 years. GP specialists admitted fewer of their patients, and GPs with low out-of-hours experience admitted more. Contacts with GPs in the highest quarter in prior admission proportion compared with the lowest quarter were associated with a substantially higher admission aOR of 1.85 (95% CI 1.74–1.96) for patients aged 0–10 years, an aOR of 1.80 (95% CI 1.75–1.85) in patients aged 11–69 years, and an aOR of 1.73 (95% CI 1.66–1.82) for patients aged 70 years and older.

Critical conditions

Contacts with GPs in the highest quarter in prior admission proportion compared with the lowest quarter were associated with higher odds also for admissions for critical conditions (aOR of 1.05, 95% CI 0.68–1.64 for patients aged 0–10 years, an aOR of 1.40, 95% CI 1.26–1.55 in patients aged 11–69 years, and an aOR of 1.16, 95% CI 1.03–1.29 for patients aged 70 years and older). Male GPs referred fewer patients aged 0–10 discharged with a critical condition (aOR 0.69, 95% CI 0.51–0.92).

Thirty-day unplanned hospital use

Higher GP age and seeing a male GP were associated with slightly fewer unplanned hospital admissions and lower costs 30 days after

Table 1. Out-of-hours contacts with GPs[a] in Norway 2008–2016: characteristics of patients and GPs weighted by the number of index contacts.

	Patients 0–10 years	Patients 11–69 years	Patients 70 years and older
Patient characteristics by number of index contacts			
All	871,947	2,553,888	509,798
Mean age, years (SD)	4 (2.8)	37 (16.3)	80 (6.9)
Male (%)	470,113 (54)	1,124,544 (44.0)	211,904 (41.6)
Low education[b] (%)	—	864,538 (37.5)	233,814 (46.1)
Immigration status[c] (%)	201,602 (23.1)	397,065 (15.6)	16,924 (3.3)
Unplanned hospital admission previous month (%)	23,482 (2.7)	133,622 (5.2)	63,960 (12.6)
Elective hospital admission previous month (%)	57,215 (6.6)	298,059 (11.7)	111,782 (21.9)
Outpatient clinic visits previous month (%)	64,480 (7.4)	320,522 (12.6)	114,415 (22.4)
Regular GP visits previous month (%)	81,592 (9.4)	228,862 (9.0)	56,239 (11.0)
Charlson Comorbidity Index[d], mean (SD)	0.01 (0.11)	0.05 (0.39)	0.24 (0.90)
Unplanned hospital admission next 10 h (%)	53,790 (6.2)	317,340 (12.4)	131,552 (25.8)
Unplanned hospital admission with urgent condition next 10 h (%)	785 (0.09)	17,334 (0.68)	16,630 (3.3)
Unplanned hospital admission with hip fracture (ICD-19 S72) next 10 h (%)	—	—	4,464 (0.9)
Unplanned hospital admission next 30 days (%)	94,490 (10.9)	510,744 (20)	197,649 (38.8)
Death within 30 days	—	5,292 (0.2)	23,509 (4.6)
GP characteristics weighted by index contacts			
Contacts with male physicians, *n* (%)	670,904 (76.9)	1,950,322 (76.4)	385,178 (75.6)
GP age, mean (SD)	43.6 (9.2)	43.7 (9.3)	43.7 (9.3)
Contacts with GP specialists, *n* (%)	4,240,124 (48.6)	1,234,398 (48.3)	239,740 (47)
Contacts with GPs with low OOH[e] experience, *n* (%)	93,549 (10.7)	283,273 (11.1)	61,311 (12.0)
Physician prior admission proportion[f] %, median (IQR[g])	10.8 (6.7–15.6)	12.1 (7.0–16.3)	11.9 (7.6–16.8)

[a]General practitioner.
[b]Completed less than 13 years in school.
[c]Immigrants and Norwegian-born to immigrant parents.
[d]Based on diagnoses from the last hospital visit previous month.
[e]Out-of-hours.
[f]The proportion of out-of-hours consultations resulting in unplanned hospital admissions the previous 4 months.
[g]Interquartile range.

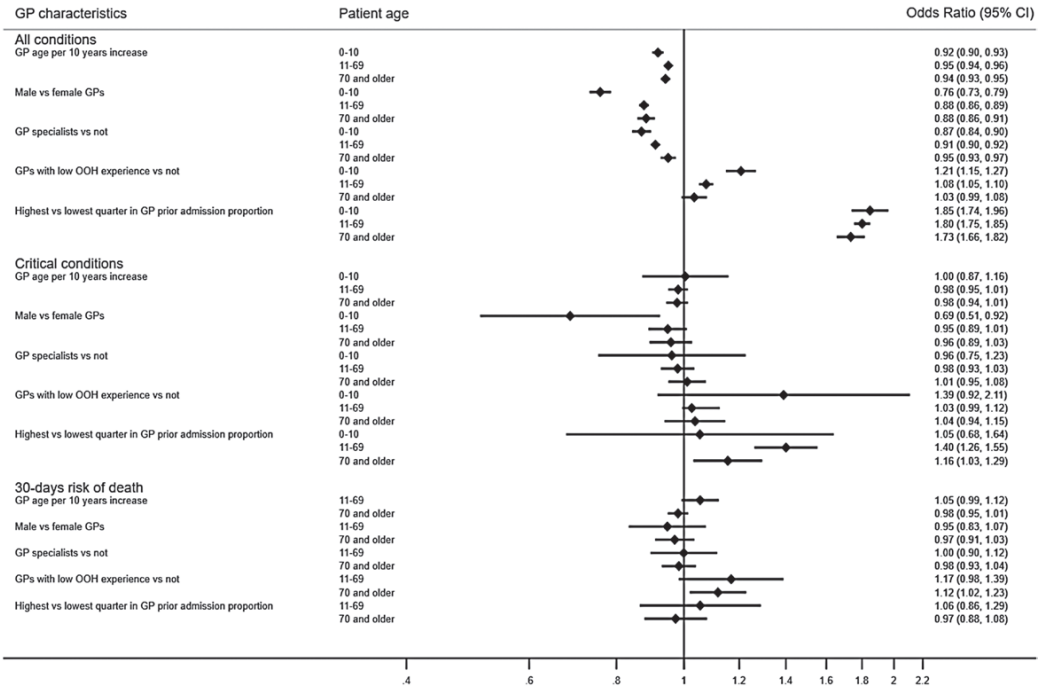


Fig. 2. All Norwegian out-of-hours contacts 2008–2016. Odds ratios for unplanned hospital admission for all conditions, critical conditions, and 30-day risk of death after the index contact. The associations were computed by comparing patients in the same 10-year age groups visiting the same out-of-hours services on the same weekdays, during the same year, month, and 8-h time unit, and were adjusted for patient sex, age, and age squared.

the index contact (Table 2). GPs' speciality status and out-of-hours experience did not substantially influence 30-day unplanned hospital admissions or costs. Contacts with GPs in the highest quarter in prior admission proportion compared with the lowest quarter were estimated to result in more unplanned hospital admissions in the 30 days following the index contact (adjusted difference per 1,000 GP contacts of 35 [95% CI 31–39] in patients 0–10 years, and 80 [95% CI 71–90] more in patients 70 years and older) and also higher 30-day unplanned hospital costs (adjusted difference per 1,000 GP contacts of 54,301€, 95% CI 36,557–72,045€ for patients aged 0–10 years, 127,344€, 95% CI 99,685–155,002€ in patients aged 11–69 years, and 412,353€, 95% CI 285,561–539,145€ for patients aged 70 years and older).

Thirty-day risk of death

There was little evidence of any associations between GP characteristics and 30-day risk of death after the index contact, however we observed 12%–17% increased odds of death associated with GPs with low out-of-hours experience (Fig. 2).

Analysis of exposure independence assumption and sensitivity analyses

After applying the matching procedure, we found weak or no associations between the patients and GP characteristics (Supplementary Table 1), supporting our assumption of independence between possible confounders and GP characteristics. The estimates of the GPs' prior admissions proportion were not substantially affected by adjustments for the other GP characteristics (Supplementary Fig. 1). The aOR for immediate admission after contact with a non-GP was slightly higher compared with contact with a GP (1.11, 95% CI 1.09–1.14 for patients 0–10 years, 1.08, 95% CI 1.06–1.09, for patients 11–69 years, and 1.02, 95% CI 1.00–1.04 for patients aged 70 years and older) (Supplementary Table 3).

Discussion

Summary

This study suggested substantial impact from GP characteristics on unplanned hospital admissions following contact with the out-of-hours services. GP age and sex showed modest associations with immediate unplanned hospital admissions. In contrast, the GPs' prior admission proportion was strongly associated with both immediate unplanned hospital admissions and 30-day unplanned hospital admissions and costs. Notably, GPs with a previously higher tendency of admitting patients also more often admitted patients with critical conditions, indicating that a more restrictive referral practice may threaten patient safety through missing out on critical cases. However, there was little evidence of the GP characteristics affecting the 30-day risk of death.

Strengths and limitations

The Norwegian out-of-hours services model with the primary care gatekeeper function resembles the systems in many western European countries providing external validity.[10] Our large study size with comprehensive register data, made it possible with close matching to avoid confounding and still achieve precise estimates

with CIs reasonably narrow. GPs have a key role in Norwegian out-of-hours services, but the recent pressure on primary care services may threaten their position in this setting. Our study did not cover non-GPs working in out-of-hours services. However, non-GP physicians working out-of-hours did not deviate substantially in immediate admissions, results in concurrence with previous research.[14]

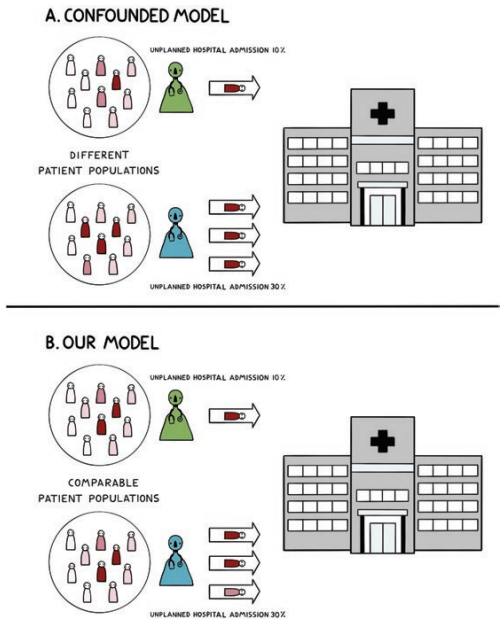


Fig. 3. (A) Different patient populations, where differences in unplanned hospital admission proportion are affected by differences in patient morbidity. (B) Assuming comparable patient populations, where differences in unplanned hospital admission proportion better reflect the differences between the GPs' decisions.

Confounding is the main concern when comparing clinical practice between physicians. A commonly used approach is to use multivariable adjustment to control for confounding. However, this requires detailed information on all important confounders, as well as no measurement error. This is an assumption we find hard to justify. Therefore, we designed our study to mimic the situation of a random distribution of patients meeting different GPs (Fig. 3). Our sensitivity analyses supported our assumption of independence between patient and GP characteristics at the time of the contact. Nevertheless, this was an observational study, and residual confounding cannot be ruled out.

Comparison with existing literature

Many approaches have been made to disentangle the factors and mechanisms of importance for decisions on referrals and hospital admissions.[2–7,22–34] With our study design aimed to limit confounding, we found that both older and male GPs were more restrictive in their admissions, which concurs with other studies' findings.[6,7,26,31,34,35] GP specialist status and out-of-hours experience showed some associations with unplanned hospital admissions for the youngest patient group, in concordance with previous literature where specialists in general practice were found to refer fewer of their patients.[26,35] However, the associations were weaker for the older patients.

Implications for research and practice

The increasing pressure on the healthcare system, including rising healthcare expenditure particularly on specialized healthcare, is challenging. New policies often aim to reduce unnecessary hospital use, where reducing variations in unplanned hospital admissions is one of many targets. Still, there is insufficient knowledge on the consequences of this variation. In this study, we found that older and male GPs admitted fewer of their patients in all patient age groups

Table 2. Change in number of unplanned hospital admissions and the cumulative costs from hospital stays starting within 30 days following the index contact according to GP characteristics, presented per 1,000 GP contacts. Linear regression analyses of all Norwegian out-of-hours contacts 2008–2016.

Thirty-day unplanned hospital use[a]	Patients 0–10 years		Patients 11–69 years		Patients 70 years+	
	Change	95% CI	Change	95% CI	Change	95% CI
Thirty-day unplanned admissions per 1,000 patient contacts						
GP[b] age per 10 years	−5	−7 to −4	−5	−6 to −4	−8	−11 to −5
Male vs female GPs	−17	−20 to −14	−14	−16 to −12	−16	−22 to −10
GP specialists vs not	−8	−10 to −6	−11	−13 to −9	−5	−10 to 0
GPs with low OOH[c] experience vs not	12	8 to 16	8	5 to 11	6	−3 to 15
Highest vs lowest quarter in prior admission proportion[d]	35	31 to 39	58	54 to 61	80	71 to 90
Costs from unplanned admissions starting within 30 days (EURO) per 1,000 patient contacts						
GP age per 10 years	−6,284	−12,196 to −373	−11,538	−20,108 to −2,969	−23,580	−61,524 to 14,365
Male vs female GPs	−33,546	−46,739 to −20,354	−25,988	−43,405 to −8,570	−24,057	−104,685 to 56,571
GP specialists vs not	−2,353	−12,516 to 7,811	−24,390	−39,353 to −9,427	68,106	−1,014 to 137,225
GPs with low OOH[c] experience vs not	15,789	−3,555 to 35,133	24,336	−821 to 49,493	38,445	−74,995 to 151,884
Highest vs lowest quarter in prior admission proportion[d]	54,301	36,557 to 72,045	127,344	99,685 to 155,002	412,353	285,561 to 539,145

[a]General practitioner.

[b]The associations were computed by comparing patients in the same 10-year age groups visiting the same out-of-hours services on the same weekdays, during the same month and same time period of the day, and were adjusted for patient sex, age, and age squared.

[c]Out-of-hours.

[d]GP prior admission proportion the previous four calendar months.

and that these differences sustained over 30 days. However, these differences were not reflected in the 30-day risk of death. Further, the differences in 30-day costs from specialized healthcare were modest, especially for the two oldest patient groups. The recruitment and retention problems currently seen in European out-of-hours primary care, as well as in general practice, can influence the composition of GPs and other physicians staffing primary care. In Norway, the cohort of GPs is changing toward a higher share of female and young physicians.[19] Yet, according to our results, even substantial changes in GP sex and age composition will not affect costs substantially. The GPs with the highest prior admission proportion however, had higher numbers of both immediate and 30-day admissions, and substantially higher costs. This implies potential for lowering specialized healthcare expenditure through strengthening the out-of-hours services, with emphasis on optimizing the framework for decision making, rather than raising the requirements for specialist status and experience. Considering the out-of-hours settings with a lack of time, resources, and previous knowledge of the patient, deciding whether to refer a patient to the hospital is more challenging than in normal hours primary care. Facilitating opportunities to confer with a peer or a more specialized physician, implementing new technical solutions like shared patient journals, decision support, and feedback on referrals and patient outcomes may help strengthen the decisions and reduce unwanted variation.

Importantly, this study also recognizes the differences in admissions for critical conditions that may nuance the picture, suggesting that more restrictive referral practices may delay admissions for such critical urgent cases and threaten patient safety. This is an important aspect in the use of referral rates as quality measures for primary care physicians,[36] and in the debate on limiting referral options on the individual GP level as a means to reduce hospital admissions.

We found no apparent associations with short term risk of death from differences in admission practices, a result that is reassuring from a patient perspective. However, the increase in 30-day risk of death after contact with GPs with low out-of-hours experience should receive further investigations.

Conclusions

This study's results provide evidence of substantial differences between GP admission practices, indicating a need for systematic work to optimize the framework for GPs' admission decisions. However, raising the requirements for experience and specialist status, or altering the age or sex in the group of GPs staffing the out-of-hours services, may not affect the consequences of the observed differences. Improving feedback on both GP admission practices and patient outcomes in the existing out-of-hours services system are possible targets.

Supplementary material

Supplementary material is available at *Family Practice* online.

Funding

This project was funded by the Research Council of Norway with grant numbers 256579 and 295989.

Ethical approval

This study has been approved by the Regional Committees for Medical and Health Research Ethics (2016/2158/REK midt). Participant consent was not required.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. All authors have completed the ICMJE COI form.

References

1. Sripa P, Hayhoe B, Garg P, Majeed A, Greenfield G. Impact of GP gatekeeping on quality of care, and health outcomes, use, and expenditure: a systematic review. *Br J Gen Pract*. 2019;**69**(682):e294–e303.
2. Bankart MJG BR, Rashid A, *et al*. Characteristics of general practices associated with emergency admission rates to hospital: a cross-sectional study. *Emerg Med J*. 2011;**2011**(28):558–563.
3. Forrest CB, Nutting PA, von Schrader S, Rohde C, Starfield B. Primary care physician specialty referral decision making: patient, physician, and health care system determinants. *Med Decis Making*. 2006;**26**(1):76–85.
4. O'Donnell CA. Variation in GP referral rates: what can we learn from the literature? *Fam Pract*. 2000;**17**(6):462–471.
5. Calnan M, Payne S, Kemple T, Rossdale M, Ingram J. A qualitative study exploring variations in GPs' out-of-hours referrals to hospital. *Br J Gen Pract*. 2007;**57**(542):706–713.
6. Ringberg U, Fleten N, Forde OH. Examining the variation in GPs' referral practice: a cross-sectional study of GPs' reasons for referral. *Br J Gen Pract*. 2014;**64**(624):e426–e433.
7. Rossdale M, Kemple T, Payne S, Calnan M, Greenwood R. An observational study of variation in GPs' out-of-hours emergency referrals. *Br J Gen Pract*. 2007;**57**(535):152–154.
8. Sullivan CO, Omar RZ, Ambler G, Majeed A. Case-mix and variation in specialist referrals in general practice. *Br J Gen Pract*. 2005;**55**(516):529–533.
9. Blinkenberg J, Pahlavanyali S, Hetlevik Ø, Sandvik H, Hunskaar S. General practitioners' and out-of-hours doctors' role as gatekeeper in emergency admissions to somatic hospitals in Norway: registry-based observational study. *BMC Health Services Res*. 2019;**19**(1).
10. Steeman L, Uijen M, Plat E, Huibers L, Smits M, Giesen P. Out-of-hours primary care in 26 European countries: an overview of organizational models. *Fam Pract*. 2020;**37**(6):744–750.
11. Berchet C, Nader C. *The organisation of out-of-hours primary care in OECD countries*. Paris; 2016. Report No: 89.
12. Marchand C, Peckham S. Addressing the crisis of GP recruitment and retention: a systematic review. *Br J Gen Pract*. 2017;**67**(657):e227–e237.
13. Svedahl ER, Pape K, Toch-Marquardt M, Skarshaug LJ, Kaspersen SL, Bjorngaard JH, *et al*. Increasing workload in Norwegian general practice - a qualitative study. *BMC Fam Pract*. 2019;**20**(1):68.
14. Sandvik H, Hunskar S. [Working style among regular general practitioners and other doctors in the out-of-hours services]. *Tidsskr Nor Laegeforen*. 2010;**130**(2):135–138.
15. The Norwegian Directorate of Health. KUHR-databasen 2019. https://www.helsedirektoratet.no/tema/statistikk-registre-og-rapporter/helsedata-og-helseregistre/kuhr.
16. Bakken IJ, Ariansen AMS, Knudsen GP, Johansen KI, Vollset SE. The Norwegian Patient Registry and the Norwegian Registry for Primary Health Care: Research potential of two nationwide health-care registries. *Scand J Public Health*. 2020;**48**(1):49–55.
17. Statistics Norway. Statistisk sentralbyrå. https://www.ssb.no/en.
18. Norwegian Insitute of Public Health. The Norwegian Cause of Death Registry 2020. https://www.fhi.no/en/hn/health-registries/cause-of-death-registry/.
19. The Norwegian Directorate of eHealth. The regular general practitioner register (Fastlegeregisteret) 2020. https://helsedata.no/forvaltere/helsedirektoratet/fastlegeregisteret/.

20. The Norwegian Diractorate of Health. Activity based funding 2016 (Innsatsstyrt finansiering 2016). 12/2015. Report No. ISBN-nr. 978-82-8081-417-3.
21. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;**40**(5):373–383.
22. Evans E AH, Edwards A. Reducing variation in general practitioner referral rates through clinical engagement and peer review of referrals: a service improvement project. *Qual Prim Care* 2011;**19**(4):263–272.
23. Barnett ML, Song Z, Landon BE. Trends in physician referrals in the United States, 1999-2009. *Arch Intern Med*. 2012;**172**(2):163–170.
24. Burton CD, McLernon DJ, Lee AJ, Murchie P. Distinguishing variation in referral accuracy from referral threshold: analysis of a national dataset of referrals for suspected cancer. *BMJ open*. 2017;**7**(8):e016439.
25. Reid FD, Cook DG, Majeed A. Explaining variation in hospital admission rates between general practices: cross sectional study. *BMJ*. 1999;**319**(7202):98–103.
26. Franks P, Williams GC, Zwanziger J, Mooney C, Sorbero M. Why do physicians vary so widely in their referral rates? *J Gen Intern Med*. 2000;**15**(3):163–8.
27. Geissler KH. Differences in referral patterns for rural primary care physicians from 2005 to 2016. *Health Serv Res*. 2020;**55**(1):94–102.
28. Grondahl JR, Fossdal O, Hauge-Iversen T, Husebye E, Rosvold EO, Kongshavn T. [Admissions to the medical department - who admits and why]. *Tidsskr Nor Laegeforen*. 2018;**138**(8).
29. Gunther S, Taub N, Rogers S, Baker R. What aspects of primary care predict emergency admission rates? A cross sectional study. *BMC Health Serv Res*. 2013;**13**:11.
30. Huntley A, Lasserson D, Wye L, Morris R, Checkland K, England H, *et al*. Which features of primary care affect unscheduled secondary care use? A systematic review. *BMJ Open*. 2014;**4**(5):e004746.
31. Ingram JC, Calnan MW, Greenwood RJ, Kemple T, Payne S, Rossdale M. Risk taking in general practice: GP out-of-hours referrals to hospital. *Br J Gen Pract*. 2009;**59**(558):e16–e24.
32. McBride D, Hardoon S, Walters K, Gilmour S, Raine R. Explaining variation in referral from primary to secondary care: cohort study. *BMJ*. 2010;**341**:c6267.
33. Oslislo S, Heintze C, Schmiedhofer M, Möckel M, Schenk L, Holzinger F. How to decide adequately? Qualitative study of GPs' view on decision-making in self-referred and physician-referred emergency department consultations in Berlin, Germany. *BMJ open*. 2019;**9**(4):e026786.
34. Gehring ND, Kebbe M, Rathwell S, Perez A, Peng C, Zendher E, *et al*. Physician-related predictors of referral for multidisciplinary paediatric obesity management: a population-based study. *Fam Pract*. 2021.
35. Ringberg U, Fleten N, Deraas TS, Hasvold T, Forde O. High referral rates to secondary care by general practitioners in Norway are associated with GPs' gender and specialist qualifications in family medicine, a study of 4350 consultations. *BMC Health Serv Res*. 2013;**13**:147.
36. Love T, Dowell AC, Salmond C, Crampton P. Quality indicators and variation in primary care: modelling GP referral patterns. *Fam Pract*. 2004;**21**(2):160–165.

Supplementary material Paper II

Supplementary material

I) Data extraction and organization

This study is based on a large data set linking public healthcare registers in Norway to demographic information about the patients and characteristics of the physicians. It comprises all contacts with the primary and secondary healthcare services in Norway from 2008-2016, except contacts with, and admissions to psychiatric hospitals.

All patient contacts with Norwegian publicly funded general practitioner services generate a unique claim for reimbursement. These claims are submitted from the treating physician or service to the Norwegian Health Economics Administration and are registered in the Control and Payment of Health Reimbursement Register (KUHR).(13) For this study, we used the reimbursement claims data from all out-of-hours contacts in Norway from 2008-16, preformed between 16:00 and 08:00 on weekdays, or during both day or night on Saturday, Sunday and public holidays. These claims include patient ID, time, patient diagnoses, a unique physician identification number, and the type of contact, (e.g. telephone, consultation, home-visits, including a specific code for out-of-hours work. We selected all claims containing the specific codes for out-of-hours work where physicians assessed patients, face-to-face or by telephone (codes: 2ak, 2fk, 1ak, 1bk, 1g). We excluded claims containing codes for home visits (codes: 2nk, 11ak, 11nk, 21k). We included claims only from physicians also working as regular General Practitioners (GPs) during office hours. This allowed us to obtain complementary information about their characteristics such as age, sex, and specialist status from the Norwegian general practitioner register. We excluded claims from contacts where the patient's regular GP were present at the out-of-hours services, since such contacts could have been arranged between the patient and the GP.

II) Study design

Our assumptions are based on the following:

Patients usually do not choose when to get acutely ill, and we assume that in most cases they have no knowledge of which physicians are on-call in their area. In the Norwegian out-of-hours services, there is often only one physician serving an area per time. (However, there is a trend towards centralizing the out-of-hours services, resulting in larger out-of-hours services, with several physicians on-call at the same time). Still, there may be local systematic differences in which physicians are on call, e.g. physicians with a family avoiding work in the week-ends, or on public holidays, and some physicians avoiding night shifts. We know that there are differences in morbidity among patients visiting the health services at different times, e.g. higher patient morbidity during night time. There may also be more specialists in some areas, and thus serving a selected population. Such differences in which physicians are on call could affect the associations, if they are not handled right. That is, local variation in patient severity visiting out-of-hours services could be associated with the characteristics of the physicians on call. When designing this study, we aimed to find groups of comparable patients, for which the only systematic difference being which physician they met at the out-of-hours service. Our basic assumption is that patients in similar age groups may be comparable within the same weekday in the same month and year in the same out-of-hours service. Hence, we matched patients in groups where we assumed that the characteristics of the physician(s) on-call were independent of the characteristics of the patients. Each matched group combines the following characteristics of the patients:

1) Patients being in same 10-year age groups

2) Patients visiting the same out-of-hours service

In Norway, all municipalities have a unique municipality code. The extensive collaboration between municipalities leads to shifting geographical locations of many out-of-hours services, especially in the scarcely populated municipalities. The KUHR claims lack information about the actual geographical location where the contact took place. However, the physicians will mainly work within only one out-of-hours service at the time, although the same out-of-hours service may often serve the population from several municipalities.

We solved this problem by linking the unique patient ID from each claim to the municipality code where patients were registered (Statistics Norway). Although some patients fall ill when traveling, we assume that patients visit the out-of-hours services in their home area in the major part of the cases. To define the location of the out-of-hours services, we used the municipality codes´ modal value among patients within physicians per week. As a result, the municipality where the majority of the patients (within one week, seeing the same physician) live was defined as the municipality where the contacts for this physician took place that particular week. Thus, this could change through the time periods in the data material. If the value was the same for two or more municipalities, we chose the lower code, as this often represents a larger municipality.

3) Patients visiting the out-of-hours service the same weekday in the same month and year

This restriction limits the possible effect of differences in local area healthcare demands during the week (e.g. higher demand on Monday and Friday evenings) and in morbidity throughout the seasons (e.g. influenza epidemic during wintertime), and the possible time trends over the years. We defined *public holidays* as Sundays, as the regular GPs´ offices are closed on these days.

4) Patients visiting the out-of-hours service in the same time unit (8-hours partitions) during the day.

Based on the assumption that there can be systematic differences between daytime, afternoon and nighttime in both patients visiting the out-of-hours-services, and the organization of out-of-hours services, we chose to divide 24 hours into three time units; 16.00-23.59; 00.00-08.00 and 08.00-16.00 on weekends.

By matching patients in groups and analysing only within-group variability, we effectively controlled for all confounding that was constant within each group. Our main assumption is that GP characteristics were not associated with the characteristics of patients' visits on similar weekdays in the same out-of-hours service, year, months, time of day and age group. As an example, we compare patients in a given 10-year age group, visiting a particular out-of-hours service on a Tuesday evening in January 2015 with patients in the same 10-year age

group visiting the same out-of-hours service one of the three other Tuesday evenings that particular month in 2015.

GPs on-call

For about 70 % of patients, there was only one GP available within each 8-hour time period, and each 10-year age group of patients. We could not rule out any selection of patients between GPs if two or more GPs were on-call at the same time, e.g. higher probability of a female patient seeing a female GP, or the more experienced GP seeing the patient with a more urgent condition. To handle this problem and avoid the effect of patient selection among GPs, we used the weighted average of GP characteristics within each 8-hour time unit in each service using number of contacts as weights. This means, for a given 8-hour unit, we used the probability of seeing, e.g. a female GP, based on the share of contacts with female GPs in the particular 8-hour time unit, regardless of which GP that provided the contact. Therefore, all patients within each matched group were given the same exposure values (GP characteristics).

Patient age groups

Furthermore, we chose to perform separate analyses for three age groups of patients; 0-10 years, 11-69 years, and above 70 years of age. We hypothesized that the youngest and oldest patient groups often use the out-of-hours services for acute health conditions where the clinical choice of immediate hospital admission is likely to be particularly important.

30-days costs from hospital admissions

All unplanned hospital admissions starting within 30 days after each index contacts were measured, also including stays lasting for more than 30 days. The costs from these admissions were calculated from the diagnosis-related group points (DRG) in the Norwegian patient registry, and presented in Euros. For the purposes of hospital funding, all similar hospital episodes are grouped into one of approximately 900 diagnoses-related groups. Each group is reimbursed a specific price, a DRG-point. One DRG-point valued at € 5075 based on the 2016 unit price (18) and average Euro exchange rate 2008-2016. Both admissions and costs were presented per 1000 contacts, to show the consequences in a greater context.

III) List of discharge ICD-10 diagnoses used to define "critical conditions"

S06 Intracranial injury (excluding S06.0 Commotio cerebri)

S72 Fracture of femur

I21 Acute myocardial infarction, unspecified

I22 Subsequent myocardial infarction

I26 Pulmonary embolism with mention of acute cor pulmonale

I63 Cerebral infarction

A41 Other Sepsis

K35 Acute appendicitis

K80 Cholelithiasis with cholecystitis (excluding S80.2, S80.5 and S80.8 which do not include cholecystitis.)

K56 Paralytic ileus and intestinal obstruction without hernia

<u>This list is made based on the following criteria:</u>

There is a professional consensus that these conditions, as the main rule, need treatment in hospitals (in the Norwegian healthcare model).

If a patient presents with one of these conditions in an out-of-hours service, it will most likely have negative consequences not being admitted to the hospital.

The conditions must not necessarily be easy to recognize when presented in the out-of-hours services, but it should lead to hospitalization if they are recognized.

IV) Additional analyses

Balance tests/confounder analyses

To justify our assumption of random assignment of GPs to patients within our defined clusters, we performed balance tests where we calculated the associations between the patient characteristics and the GP characteristics. These are shown in Supplementary Table 1. We made variables for the healthcare use prior to the out-of-hours contact for all patients. Together with variables for patient sex, age, education level and immigration status, these variables were used in analyses to test whether patient characteristics affected which GP they saw when visiting the out-of-hours services. As we can see from the results from Supplementary Table 1, there was little or no evidence of any differences between GP characteristics and possible patient confounders, given our design. Hence, these results provide a strong support of the confounder independence assumption of our design.

Table S1. Balance test of potential confounders[1]

	Patients 0-10 years		Patients 11-69 years		Patients 70 years+	
	Beta	**95% CI**	**Beta**	**95% CI**	**Beta**	**95% CI**
Patient age						
GP age per 10 years increase	**-0,01**	-0,02 - 0,01	**0,00**	-0,01 - 0,01	**-0,02**	-0,03 - 0,00
Male vs female GP	**-0,01**	-0,03 - 0,02	**0,00**	-0,02 - 0,01	**-0,02**	-0,05 - 0,02
GP specialists vs not	**-0,01**	-0,03 - 0,01	**0,00**	-0,01 - 0,02	**-0,04**	-0,07 - 0,00
GPs with low OOH[2] experience vs not	**-0,01**	-0,04 - 0,02	**0,01**	-0,01 - 0,03	**0,06**	0,01 - 0,11
10% increase in GP prior admission proportion[3]	**0,01**	-0,01 - 0,02	**0,00**	-0,01 - 0,01	**0,00**	-0,03 - 0,03
	OR	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
Female patients						
GP age per 10 years increase	**1,01**	1,00 - 1,02	**1,00**	1,00 - 1,01	**1,00**	0,99 - 1,01
Male vs female GP	**1,00**	0,99 - 1,02	**0,98**	0,97 - 0,99	**1,00**	0,98 - 1,03
GP specialists vs not	**1,01**	0,99 - 1,02	**0,99**	0,98 - 1,00	**0,99**	0,97 - 1,01
GPs with low OOH[2] experience vs not	**0,99**	0,97 - 1,02	**1,00**	0,99 - 1,02	**0,98**	0,95 - 1,02
10% increase in GP prior admission proportion[3]	**0,99**	0,98 - 1,01	**1,00**	0,99 - 1,01	**1,00**	0,98 - 1,02
Patient education[4]						
GP age per 10 years increase			**1,00**	1,00 - 1,01	**1,00**	0,98 - 1,01
Male vs female GP			**0,99**	0,98 - 1,00	**1,00**	0,98 - 1,03
GP specialists vs not			**0,99**	0,98 - 1,00	**0,99**	0,96 - 1,01
GPs with low OOH[2] experience vs not			**1,02**	1,00 - 1,03	**1,01**	0,97 - 1,05
10% increase in GP prior admission proportion[3]			**1,00**	0,99 - 1,01	**0,99**	0,97 - 1,01

Patient immigration status[5]								
GP age per 10 years increase	**1,00**	0,99 - 1,01		**1,00**	1,00 - 1,01		**1,00**	0,96 - 1,03
Male vs female GP	**1,01**	0,99 - 1,03		**0,99**	0,98 - 1,00		**1,00**	0,94 - 1,07
GP specialists vs not	**1,01**	1,00 - 1,03		**1,01**	1,00 - 1,02		**0,99**	0,93 - 1,04
GPs with low OOH[2] experience vs not	**1,01**	0,98 - 1,04		**1,00**	0,98 - 1,03		**1,03**	0,94 - 1,14
10% increase in GP prior admission proportion[3]	**0,99**	0,98 - 1,01		**0,99**	0,98 - 1,01		**0,96**	0,91 - 1,02
Patients with unplanned hospital admission previous month								
GP age per 10 years increase	**1,02**	0,99 - 1,05		**1,00**	0,99 - 1,01		**1,02**	1,00 - 1,04
Male vs female GP	**0,98**	0,93 - 1,03		**0,98**	0,96 - 1,01		**0,99**	0,95 - 1,03
GP specialists vs not	**1,01**	0,97 - 1,06		**0,99**	0,97 - 1,01		**0,99**	0,96 - 1,03
GPs with low OOH[2] experience vs not	**0,99**	0,92 - 1,07		**1,04**	1,01 - 1,08		**1,05**	0,99 - 1,10
10% increase in GP prior admission proportion[3]	**1,01**	0,97 - 1,05		**1,01**	0,99 - 1,03		**1,00**	0,97 - 1,03
Patients with elective hospital admission previous month								
GP age per 10 years increase	**0,99**	0,98 - 1,01		**1,00**	0,99 - 1,01		**1,00**	0,99 - 1,02
Male vs female GP	**1,00**	0,96 - 1,03		**0,98**	0,97 - 1,00		**1,00**	0,97 - 1,03
GP specialists vs not	**0,97**	0,94 - 1,00		**0,99**	0,98 - 1,00		**1,02**	1,00 - 1,05
GPs with low OOH[2] experience vs not	**0,99**	0,94 - 1,04		**1,00**	0,97 - 1,02		**1,01**	0,96 - 1,05
10% increase in GP prior admission proportion[3]	**0,99**	0,97 - 1,02		**1,01**	1,00 - 1,02		**1,01**	0,98 - 1,03
Patients with out-patient clinic visits previous month								
GP age per 10 years increase	**1,00**	0,98 - 1,01		**1,00**	0,99 - 1,01		**1,00**	0,99 - 1,02
Male vs female GP	**0,99**	0,96 - 1,03		**0,98**	0,97 - 1,00		**1,01**	0,98 - 1,04
GP specialists vs not	**0,97**	0,94 - 0,99		**0,99**	0,98 - 1,00		**1,02**	0,99 - 1,04
GPs with low OOH[2] experience vs not	**1,00**	0,95 - 1,05		**1,00**	0,98 - 1,02		**1,00**	0,96 - 1,05
10% increase in GP prior admission proportion[3]	**0,99**	0,97 - 1,02		**1,01**	1,00 - 1,02		**1,01**	0,99 - 1,04
Patients with GP visits previous month								
GP age per 10 years increase	**1,01**	0,99 - 1,02		**1,01**	1,00 - 1,02		**1,01**	0,99 - 1,03
Male vs female GP	**1,02**	0,99 - 1,05		**1,01**	0,99 - 1,03		**1,02**	0,98 - 1,06
GP specialists vs not	**1,02**	0,99 - 1,04		**1,00**	0,98 - 1,01		**1,02**	0,98 - 1,06
GPs with low OOH[2] experience vs not	**0,97**	0,93 - 1,02		**1,01**	0,98 - 1,04		**1,03**	0,97 - 1,09
10% increase in GP prior admission proportion[3]	**0,99**	0,97 - 1,02		**0,99**	0,98 - 1,01		**0,99**	0,96 - 1,02
	Beta	**95% CI**		**Beta**	**95% CI**		**Beta**	**95% CI**
Charlson Comorbidity Index previous month [6]								
GP age per 10 years increase	**0,00**	0,00 - 0,00		**0,00**	0,00 - 0,00		**0,00**	0,00 - 0,01
Male vs female GPs	**0,00**	0,00 - 0,00		**0,00**	0,00 - 0,00		**0,00**	-0,01 - 0,01
GP specialists vs not	**0,00**	0,00 - 0,00		**0,00**	0,00 - 0,00		**0,00**	-0,01 - 0,01
GPs with low OOH experience vs not	**0,00**	0,00 - 0,00		**0,00**	0,00 - 0,00		**0,00**	-0,02 - 0,02
10% increase in GP prior admission proportion	**0,00**	0,00 - 0,00		**0,00**	0,00 - 0,00		**0,00**	-0,01 - 0,01

[1] The associations were computed by comparing patients in the same ten-year age groups visiting the same out-of-hours services on the same weekdays, during the same month and same time period of the day, and were adjusted for patient sex, age and age squared.
[2] Out-of-hours
[3] On a continuous scale
[4] Completed less than 13 years in school
[5] Immigrants and Norwegian-born to immigrant parents
[6] Based on diagnoses from the last hospital visit in the previous month

Sensitivity analyses I

We made sensitivity analyses for all primary outcomes, where we adjusted all exposure variables for each other. This did not affect the results, and in particular, the strong effect of the GP prior admission proportion were largely unchanged. See Supplementary Figure 1.

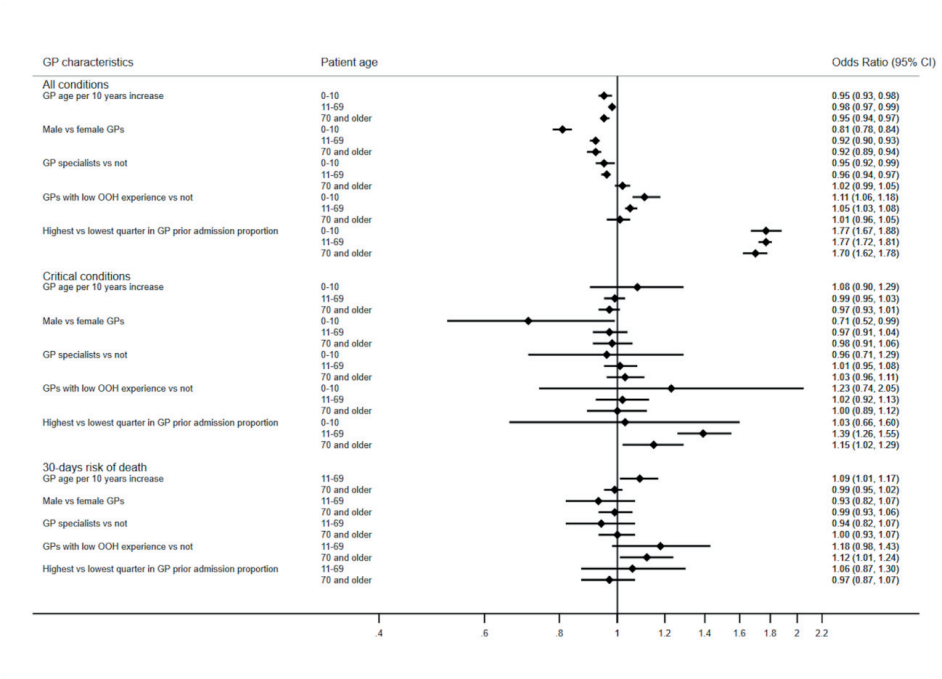


Figure S1: All Norwegian out-of-hours contacts 2008-2016. Odds ratios for unplanned hospital admission for all conditions, critical conditions and 30-days risk of death after the index contact, all exposure variables adjusted for each other.
[1] The associations were computed by comparing patients in the same 10-year age groups visiting the same out-of-hours services on the same weekdays, during the same year, month and 8-hour time unit, and were adjusted for GP characteristics, patient sex, age and age squared.

Sensitivity analyses II

As a sensitivity analysis we wished to test for a health condition where we assumed there would be no difference in unplanned hospital admissions between the GPs. Hip fracture is a relatively common and recognizable condition among elderly patients, with a consensus of requiring immediate hospital admission. Hence, we defined the outcome variable 'control

condition' as admissions with the discharge diagnosis of hip fracture (S72 in ICD-10) in patients 70 years and older. For patients discharged with the control condition (hip fracture), we found no associations between any of the GP characteristics and unplanned hospital admission, supporting our model. See Supplementary Table 2.

Table S2. Odds ratios for unplanned hospital admission resulting in a discharge diagnosis of S72 Hip fracture after index consultation for patients 70 years and older[1]

	OR	95% CI
Control condition (Hip fracture, ICD-10 S72)		
GP age per 10 years increase	**1,03**	0,96 - 1,10
Male vs female GPs	**1,00**	0,99 - 1,01
GP specialists vs not	**0,99**	0,98 - 1,01
GPs with low OOH[2] experience vs not	**1,00**	0,98 - 1,02
10 % increase in GP prior admission proportion[3]	**0,98**	0,88 - 1,10
Highest vs lowest quarter in prior admission proportion[4]	**1,07**	0,80 1,18

[1] The associations were computed by comparing patients in the same ten-year age groups visiting the same out-of-hours services on the same weekdays, during the same month and same time period of the day, and were adjusted for patient sex, age and age squared.
[2] Out-of-hours
[3] GP prior admission proportion the previous four calendar months, on a continuous scale.

Sensitivity analyses III

In this study, we chose to include only patient contacts with active GPs working in the out-of-hours setting, as the pressure in the healthcare services may threaten the contribution from GPs in the out-of-hours services. Also, non-GPs staffing the out-of-hours services comprise a heterogenous group of physicians (e.g. hospital physicians in various specialities, interns, PhD fellows and locums).

This further enabled us more comprehensive details on the physicians included in the study, (i.e. age and specialist status) from the General Practitioner Register.

Even if studying differences between the performance of GPs and non-GPs was not in the scope of this study, we performed sensitivity analyses to ensure to detect any striking differences in the referral practices between the groups. The analyses showed a slightly higher OR for immediate unplanned admission after contact with a non-GP compared to a GP, as shown in Supplementary Table 3.

Table S3. Odds ratios for unplanned hospital admission after an out-of-hours contact with a non GP compared to a GP.

Table S3

Patient age group	Patients 0-10 years		Patients 11-69 years		Patients 70 years+	
	OR	95% CI	OR	95% CI	OR	95% CI
All conditions						
Non-GPs vs GPs	**1,11**	1,09 - 1,14	**1,08**	1,06 - 1,09	**1,02**	1,00 - 1,04

Sensitivity analyses IV

In our material, some GPs were registered with a very high number of claims (> 10 000 contacts over two years or >435 contacts per month, representing the 98 percentile), meaning they had either worked significantly more than the average or that this represents wrong registration. In both cases, this could affect our results. To deal with this, we made sensitivity analyses where we excluded the GPs registered with > 10 000 claims per two years, and/or 435 contacts per month. This did not affect the results (data not shown).

Sensitivity analyses V

Table S4. All Norwegian out-of-hours contacts 2008-2016. Odds ratios for unplanned hospital admission for all conditions after the index contact, presented for each 10-years age group.

Table S4

Patient age group	0-10		11-19		20-29		30-39		40-49		50-59		60-69		70-79		80-89		90 years and older	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
All conditions																				
GP age per 10 years increase	**0,92**	0,90 - 0,93	**0,97**	0,95 - 0,99	**0,96**	0,94 - 0,98	**0,94**	0,92 - 0,96	**0,94**	0,92 - 0,96	**0,94**	0,93 - 0,96	**0,94**	0,93 - 0,96	**0,94**	0,93 - 0,96	**0,94**	0,92 - 0,96	**0,94**	0,89 - 0,98
Male vs female GPs	**0,76**	0,73 - 0,79	**0,90**	0,87 - 0,94	**0,86**	0,83 - 0,89	**0,87**	0,83 - 0,90	**0,85**	0,82 - 0,89	**0,89**	0,85 - 0,92	**0,89**	0,86 - 0,93	**0,86**	0,83 - 0,90	**0,90**	0,86 - 0,93	**0,94**	0,85 - 1,03
GP specialists vs not	**0,87**	0,84 - 0,90	**0,91**	0,88 - 0,94	**0,91**	0,88 - 0,94	**0,90**	0,87 - 0,93	**0,85**	0,82 - 0,89	**0,90**	0,87 - 0,93	**0,93**	0,90 - 0,97	**0,95**	0,92 - 0,99	**0,93**	0,90 - 0,97	**1,00**	0,92 - 1,10
GPs with low OOH experience vs not	**1,21**	1,15 - 1,27	**1,06**	1,00 - 1,13	**1,09**	1,04 - 1,15	**1,13**	1,06 - 1,19	**1,06**	1,01 - 1,13	**1,10**	1,04 - 1,16	**1,01**	0,96 - 1,07	**1,09**	1,03 - 1,16	**0,98**	0,92 - 1,04	**1,03**	0,90 - 1,19
Q4 vs Q1 in GP prior admission proportion	**1,85**	1,74 - 1,96	**1,78**	1,67 - 1,90	**1,81**	1,70 - 1,92	**1,97**	1,84 - 2,11	**1,80**	1,69 - 1,92	**1,80**	1,68 - 1,92	**1,68**	1,58 - 1,79	**1,75**	1,64 - 1,88	**1,68**	1,57 - 1,80	**1,96**	1,65 - 2,33

Adjusted odds ratio for unplanned hospital admission within 10 hours, presented for each 10-years age group.

Paper III

Out-of-hours referral to hospital – The impact of altering the referral threshold on patient safety and further health service use

Ellen Rabben Svedahl, MD, PhD candidate[1*], Kristine Pape, MD, associate professor[1], Bjarne Austad, MD, associate professor[1], Gunnhild Åberge Vie, MD, associate professor[1], Kjartan Sarheim Anthun, PhD[1,2], Fredrik Carlsen, professor[3], Neil Martin Davies[4,5], Johan Håkon Bjørngaard, professor[1,6]

1 Department of Public Health and Nursing, Norwegian University of Science and Technology, Trondheim, Norway;
2 Department of Health Research, SINTEF Digital, Trondheim, Norway;
3 Department of Economics, Norwegian University of Science and Technology, Trondheim, Norway
4 Medical Research Council Integrative Epidemiology Unit, University of Bristol, BS8 2BN, United Kingdom
5 Population Health Sciences, Bristol Medical School, University of Bristol, Barley House, Oakfield Grove, Bristol, BS8 2BN, United Kingdom
6 Nord University, Faculty of Nursing and Health Sciences, Levanger, Norway

*Corresponding author:
Ellen Rabben Svedahl MD
PhD Fellow,
Department of Public Health and Nursing
Norwegian University of Science and Technology
Post box 8905
7491 Trondheim

Phone: +47 980 80 173
E-mail: ellen.r.svedahl@ntnu.no

Keywords:
Out-of-hours services, Hospital admissions, Referrals, Patient safety, Primary care, after-hours care, general practitioners, health services research, patient admission, patient safety, referral and consultation

Abstract

Objectives

To determine the impact from altering referral thresholds from the out-of-hours services on the patients´ further use of health services and risk of death.

Design

Register-based observational study using patient data from primary and specialised health services, demographical data from Statistics Norway and the Norwegian Cause of Death Registry.

Setting

Norway

Participants

922,796 patients aged 65 years and older contacting the Norwegian out-of-hours services between 2008-2016.

Main outcome measures

Multivariable adjusted and instrumental variable associations between referrals to hospital from out-of-hours services, for further health services use and death for up to six months. Physicians' proportions of acute referrals of other patients from out-of-hours work were used as an instrumental variable (physician referral preference) for their threshold of referral in cases of clinical doubt.

Results

For elderly patients, whose referrals were attributable to their physicians' threshold for referral mean length of stay in hospital increased 3.25 days (95% CI 3.08-3.42) within the first ten days, compared with non-referred patients. Such referrals also increased six months use of outpatient specialist clinics and primary care physicians. Importantly, patients with referrals attributable to their physicians' threshold had a substantially reduced risk of death the first ten days (HR 0.53, 95% CI 0.29-0.92), an effect sustaining through the six months follow up period (HR 0.70, 95% CI 0.50-0.95).

Conclusions

Out-of-hours patients whose referrals are affected by physician referral threshold contribute substantially to the use of health services. However, the referral seems protective by reducing the risk of death in the first six months after the referral. Thus, raising the threshold for referral to lower pressure on overcrowded emergency departments and hospitals should not be encouraged without ensuring the accuracy of the referral decisions.

Introduction

Avoiding unnecessary hospitalisations will likely become a strong policy tool following increasing demand for health services from an ageing population. (1, 2) Referring or not referring the patient to a hospital for admission is a well-known dilemma for physicians working in primary care. Such gatekeeping may have a crucial influence on service utilization, but consequences for patient outcomes are poorly understood. (3, 4) In cases where the symptoms and findings are diffuse but still may imply severe illness, some physicians will refer and possibly contribute to increasing emergency department crowding and health services demand. Others will choose a wait-and-see approach or initiate other actions and possibly delay access to vital specialised care. As a consequence, in practice, many patients will have their decision of referral determined by the threshold for the physician assessing them. Emergency referrals from out-of-hours services are of particular interest since the out-of-hours physician has scarce or no knowledge about the patient and limited time, resources, and access to medical records, in contrast to the regular general practitioners (GPs) working daytime.

The consequences for patient safety and health service use from altering the referral threshold are mainly unknown and arguably complex to study. (5) Since randomisation would be practically challenging, slow and expensive, estimating the causal effects of delaying or avoiding hospitalisation is difficult. Observational data could help describe the outcomes of patients who are referred or not. However, patients who are referred are likely to be sicker than those not referred, and it is unclear whether it is possible to account for a sufficient set of covariates to control for all differences between individuals.

Instrumental variable (IV) analysis is a design that can, under some assumptions, estimate the causal effects of interventions in the presence of unmeasured confounding of the exposure-outcome relationship by using observational data. (6) Physician prescribing preferences have been widely used to estimate different treatments' comparative effectiveness and safety. (7-9) Physicians' preferences for referring, as indicated by their referral decisions about their previous patients, are known to predict their decisions regarding their future patients. (10) These referral preferences are unlikely to be related to

the co-morbidities of their current patient, especially for out-of-hours physicians seeing a particular patient for the first time.

We investigated health services use and deaths that occurred up to six months after the initial contact with the out-of-hours services for older patients previously unknown to the physician and who had few out-of-hours visits. We used the physicians' preferences for referral as a candidate instrumental variable to provide valuable and credible information about the consequences of hospital admissions for patients whose clinical presentations are less clear cut, thus referrals can be attributed to their physicians' referral threshold.

Materials and methods

Study setting

The Norwegian health care system is based on universal access, with only limited patient co-payments for GP services and outpatient specialist clinics and no co-payment for hospital admissions. (11) The out-of-hours service in Norway is part of the municipal primary care and organised as a GP cooperative which is the dominant model in Europe. (12) The service handle most acute illnesses outside office hours, and primary care physicians act as gatekeepers for secondary care. In most cases, patients have no influence on the choice of physician in the out-of-hours setting. Participating in the out-of-hours services is mandatory for regular GPs, but other physicians also staff the services and contribute with about half of the contacts. (13)

Study population & design

The study population included all patients of 65 years and older, assessed by an out-of-hours physician in Norway, either face-to-face at the out-of-hours service station or by telephone during the study period from 01.01.2008 to 31.12.2016. We included patients 65 years or older visiting out-of-hours services between 16:00 and 08:00 on weekdays and whole Saturday, Sunday and public holidays. (14) Using a unique identification number, we linked information on specialised health services use (psychiatric care was not included) (15), primary care physicians use, (14) demographical information from Statistics Norway, (16) and date of death from the Norwegian Cause of Death Registry. (17) Information from the Norwegian General Practitioner Register were linked to each patient contact by a unique

physician ID. (18) Please see appendix for a detailed description of data sources and study design.

In Figure 1 we list the criteria for study inclusion. To reduce the risk of including contacts where the physicians were familiar with the patients, we excluded contacts where one or more physicians on call had missing id, where patients had been assessed by the same physician previously during the data period from 2006-2016, where the patient met his or her regular GP, and from patients without a registered regular GP. Frequent attenders to out-of-hours services will most likely be known to the staff. (19) Hence, we excluded all contacts from frequent attenders (> 4 consultations per year). Patients who were registered with date of death before the date of contacts, were also excluded (n=68).

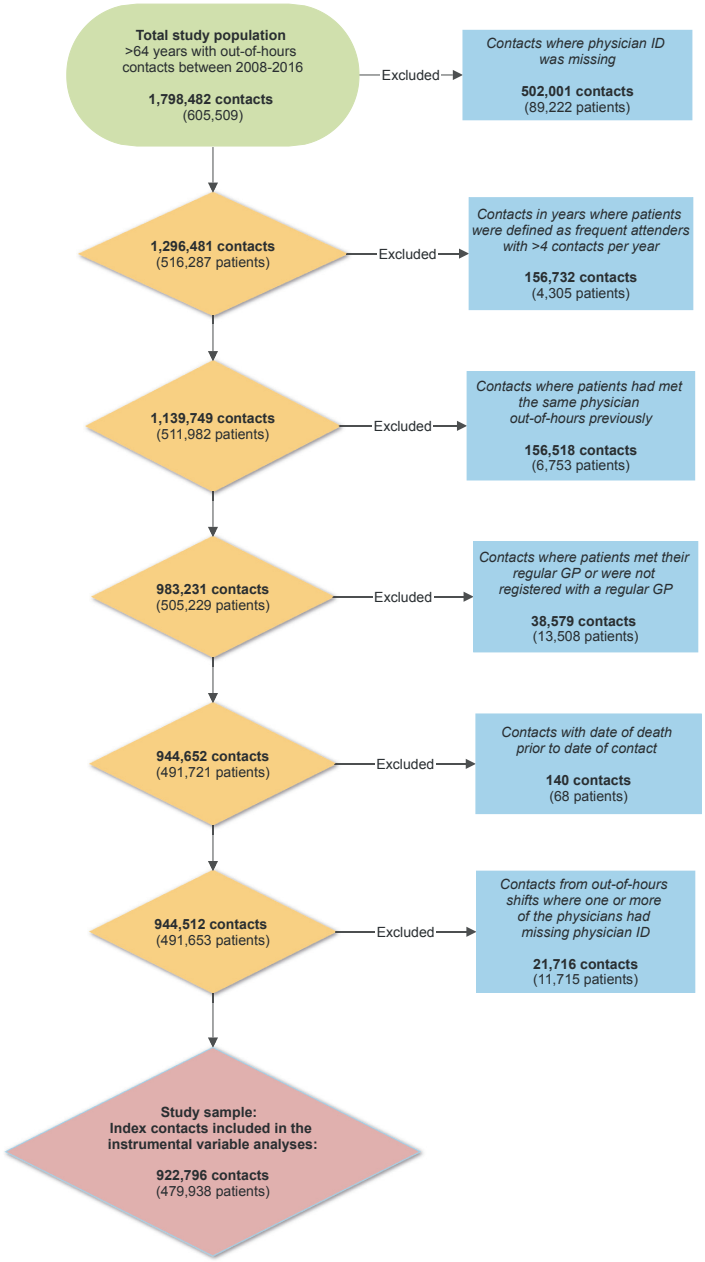


Figure 1. Flowchart

Exposures

Referral to the hospital for admission was defined as a registered unplanned visit to hospital within ten hours after an index contact with the out-of-hours services (both inpatient and outpatient visits).

Outcomes

We measured health services use as 1) hospital inpatient days (where day care treatment was counted as 0.5 days), 2) days with contacts with primary care physicians and 3) days with outpatient clinic consultations. Risk of death was measured as time to death for any reason, from the Norwegian Cause of Death registry. All outcomes were measured during four time periods after the index consultation: 0-10 days, 2) 0-30 days, 3) 0-90 days and 4) 0-180 days.

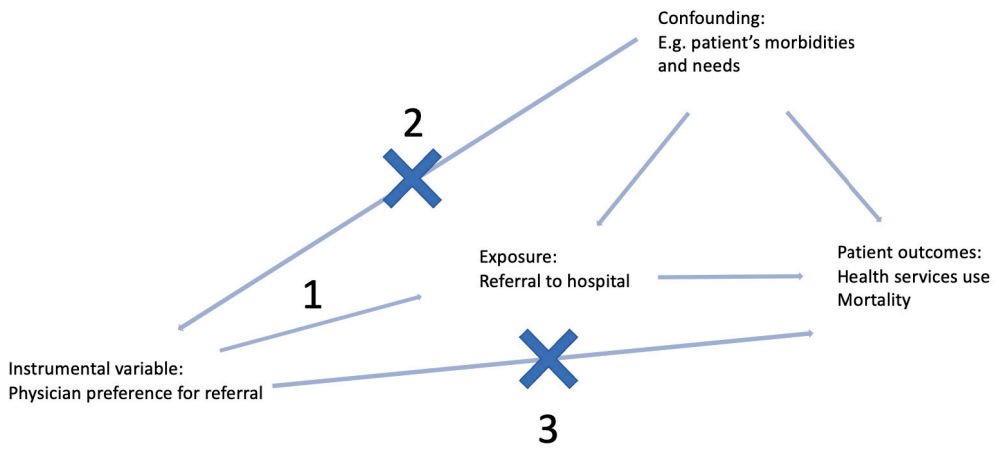


Figure 2. Directed acyclic graph showing the assumptions for using instrumental variable analyses. 1) Relevance: The instrument is associated with the exposure. 2) Independence: The instrument and the exposure have no common confounders. 3) Exclusion: The instrument does not affect the outcome other than through the exposure.

Instrumental variable analyses

We used the physician's preference for referral as a candidate instrumental variable in order to estimate the causal effect of hospital referral for patients whose refferals are attributable to their physicians referral threshold. There are three key assumptions for IV estimation: Firstly, the instrumental variable must be strongly associated with the exposure (relevance), conventionally assessed with a F-test >10 although such a dichotomization may be misleading. (20) Secondly, the instrumental variable should not be associated with any confouding factors (independence). Thirdly, the instrument must affect the outcome only through the effect on the exposure (exclusion restriction). Also a fourth assumption is relevant when assessing the effect of referrals attributable to referral threshold, namly that the instrument only affects the exposure in one direction (monotonicity). (21) We can test the first assumption by estimating the association of physicians´ referral behaviour for other patients than the index patient and the referral decision for the index patient. The second and third assumptions are not directly testable, but they can be falisified, by estimating the association of the physician referral preference and their patients morbidity and other relevant characteristics. The fourth assumption is mostly not testable, but based on subject-matter knowledge, assuming that physicians' referral threshold will most likely not have an opposite effect on a patient's chance of referral.

The instrument was calculated as the physicians' proportion of patient contacts in out-of-hours work (from 2008-2016) leading to an immediate unplanned hospital within ten hours, for male and female patients separately, assumingly through an acute referral to hospital. To detach activity beween the index contacts and the instrument, we used the physician´s admission proportion of female patients as an instrument for admitting male patients and vice versa. Still, there could be patient differences across physicians, particularly between geographical areas and time. Hence, we matched patients into groups defined by combining information on patients of the same sex, visiting the same out-of-hours station in the same 8-hour time unit during the day, within the same year. For example a woman visiting a service in the afternoon in 2015, were only compared with other women visiting the same service the other afternoons (16:00-23:59) in 2015. To avoid the effect of possible patient selection in situations where two or more GPs were on-call at the same time, we used the

weighted average of GP characteristics within each 8-hour time unit in each service. (Details of the study design are presented in Supplementary).

Statistical analyses and analytical design

We compared health service use among patients with and without referral to hospital following out-of-hours consultations, by using both multivariable adjusted analyses, and instrumental variable analyses. For the multivariable adjusted analyses we used multivariable adjusted linear regression models to investigate the associations between referral and the further health service use. We adjusted for year, month, weekday, consultation hour, patient sex, age and age squared. We used Cox regression with time from index consulation as the time axis to investigate the association between referral and mortality.

Follow-up began on the day of each index contact with the out-of-hours services and lasted for the entire follow-up period (10, 30, 90 and 180 days). We excluded contacts from the last months in 2016, where the patients could not be observed through the defined follow-up time (from July to December in the analyses for 180 days, October to December for 90 days, December for 30 and 10 days).

For the instrumental variable analysis, we used within-group-estimator for instrumental variable regression (ivreghdfe in STATA) to study the effects on further use of health services. (22) We adjusted for month, weekday, patient age and age squared, in addition to the within group estimation. Mortality was analysed with a within-matched-group estimator using stratified Cox analyses, and a two-sample IV estimator (the delta method). (23) Precision was evaluated with 95% confidence intervals (CI) with robust standard errors clustering on physician and patient ID.(22)

Assumptions and additional analyses

We assumed that our instrument was unrelated to the co-morbidities of the physicians' current patients. To support this independence assumption, we analysed associations between our instrument and possible confounding patient characteristics such as, age, sex, immigration status (yes/no), education (0-10 years vs 13 years or more), and health service

contacts 30 days before the index contact (i.e., GP visits, planned and unplanned hospital admissions, and outpatient visits), discharge diagnoses from previous hospital stays divided in main chapters (ICD10, Chapter IX Diseases of the circulatory system, and Chapter II Neoplasms). Results for these balance test of confounders are presented in Supplementary Table S1.

As sensitivity analyses for the instrumental variable analyses, we used an alternative definition of the instrument for the physician referral preference. This definition was the physicians´ proportion of patient contacts that led to a referral for all patients, excluding contacts from the same year, and the years before and after the index contact. This exclusion made the instrument less prone to effects from cases where certain physicians works certain shifts at the time of the index contact, thus seeing patients with systematically differing morbitities.

Results

Study population

During the study period, there were 1,798,482 contacts with the out-of-hours services made by 605,509 unique patients (Figure 1). After further exluding contacts not fulfilling the eligibility criteria, we were left with 922,796 contacts, from 479,938 unique patients. About one fourth of the contacts resulted in an acute referral. The mean patient age was 77.4 years, and 43% of the contacts were made by male patients (Table 1).

Table 1 Out-of-hours contacts in Norway 2008–2016:

Characteristics of the study population weighted by the number of index contacts.

Patient characteristics weighted by the number of index contacts

n= 479,938 patients, 6,752 physicians	**Contacts**	**Not referred**	**Referred**
All (%)	922,796	687,991 (74.6)	243,805 (26.4)
Mean age, years (SD)	77.4 (8.4)	77.0 (8.4)	78.5 (8.4)
Male (%)	395,765 (42.9)	289,579 (45.2)	106,186 (42.1)
Low education (%)	381,059 (41.5)	280,732 (41.0)	100,327 (42.9)
Immigration status (%)*	36,302 (3.9)	27,312 (4.0)	8,990 (3.8)
Unplanned admission to hospital previous month (%)	88,478 (9.6)	60,183 (8.8)	28,295 (12.1)
Elective contact with hospital previous month (%)	189,474 (20.5)	137,585(20.0)	51,889 (22.1)
Out-patient clinic visits previous month (planned and unplanned) (%)	192,077 (20.8)	139,687 (20.3)	52,390 (22.3)
Primary care physician visits previous month (%)	53,344 (5.8)	39,510 (5.7)	13,834 (5.9)
ICD-10 chapter I (previous hospital contact)	99,360 (10.8)	71,315 (10.4)	28,045 (12.0)
ICD-10 chapter C (previous hospital contact)	47,518 (5.15)	34,167 (5.0)	13,351 (5.7)
Physician referral preference for men, mean (SD*)	0.24 (0.1)	0.23 (0.1)	0.27 (0.1)
Physician referral preference for women, mean (SD*)	0.27 (0.1)	0.26 (0.1)	0.30 (0.1)

*Standard deviation

Use of health services

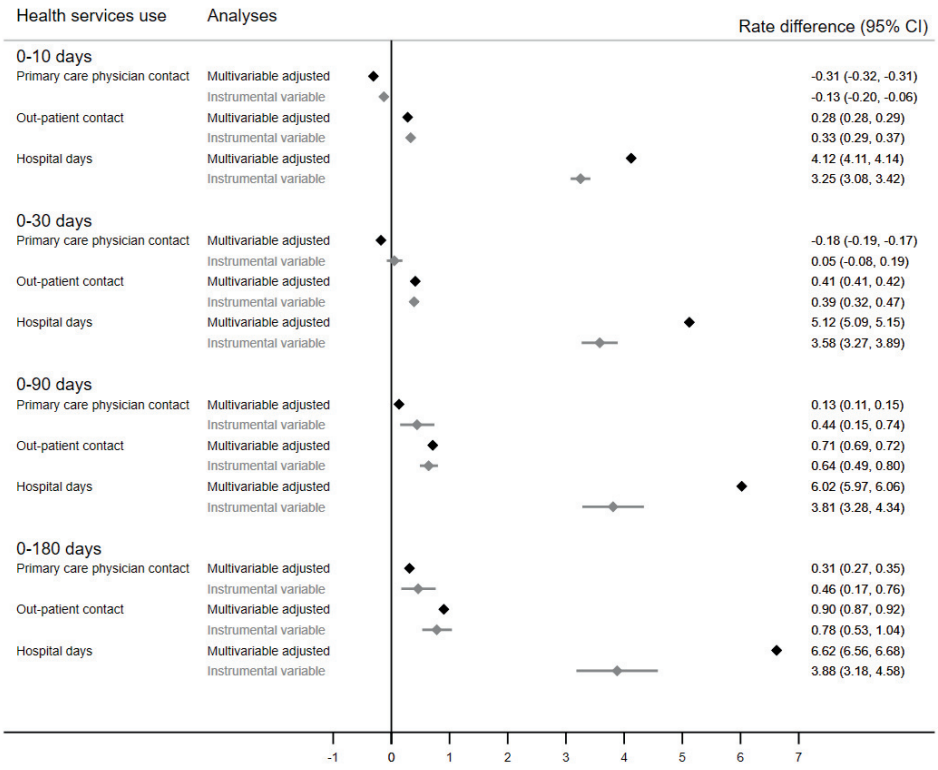


Figure 3. Mean difference between referred and non-referred patients in health service use measured as days with primary care physician contacts, days with out-patient clinic visits and days in hospital for patients aged 65 years and older, by multivariable adjusted* and instrumental variable** analyses.
*Adjusted for patient age, age squared, sex, year, month, weekday, and hour.
**Analysis within the same out-of-hours service, same time unit, year and sex. Adjusted for patient age, age squared, month, and weekday.

The instrumental variable analysis suggests an increase of 3.25 (95% CI, 3.8 to 3.42) days in hospital following an acute referral from an out-of-hours physician 0-10 days after the consultation (Figure3). This estimated effect was lower than the estimate for the multivariable adjusted analysis (p-value for difference <0.001). The substantial increase in health care use was present up to six monthts after the index consultation. There was a slight increase in days with outpatient visits, with similar estimates between the

instrumental variable and multivariable adjusted estimates. With six months follow up, there was also an increase in days with primary care physician contacts.

Table 2 Hazard ratio for death for patients 65 years and older, within 10, 30, 90 and 180 days after a referral vs no referral following contact with the out-of-hours services in Norway, 2008-2016.

Follow up after index contaxt	Multivariable adjusted*			Instrumental variable**		
	Hazard ratio	**LCI**	**UCI**	**Hazard ratio**	**LCI**	**UCI**
0-10 days	1.40	1.36	1.45	0.53	0.29	0.92
0-30 days	1.56	1.53	1.60	0.54	0.33	0.83
0-90 days	1.59	1.57	1.62	0.60	0.41	0.84
0-180 days	1.54	1.52	1.56	0.70	0.50	0.95

**Adjusted for year, month, weekday, sex, age and age^2*

***The exposure (referral) is instrumented by the physician referral preference, calculated as the share of patients of the opposite sex, who were referred in the period. Adjusted for age, age^2, month and weekday.*

Patient safety/death

In table 2 we present the hazard ratios (HR) for death within the defined follow-up times of 10, 30, 90 and 180 days. The instrumental variable estimate suggested a substantial risk reduction of referral to hospital. For the first 10 days, the risk of dying was almost halved (HR 0.53, 95% CI 0.29-0.92) suggesting a substantial protective effect of referral. This protective effect seemed to sustain, although steadily weakening through the follow-up period.

Additional analyses

There was a strong association between our instrumental variable physiscian referral preference and the probability of referral for the patient (F-value≈1,200, Table S2). Analyses of confounder balance related to the instrument showed weak or no associations in favour of our independence assumptions (Table S1). We used different definitions of the instrument, and the estimates did not substantially differ when using a different specification of the preference instrument (Table S3).

Discussion

Principal findings

In this national study with nearly 1 million out-of-hours contacts from patients aged 65 years or older, we found substantial effects of being referred to the hospital for patients whose referral was attributable to their physicians threshold for referral. Using instrumental variable analyses, we found that these patients had substantially more days in the hospital, higher use of outpatient specialist clinics and primary care physicians the following six months if they were referred. Further, while multivariable adjusted regression analyses estimated a substantially increased risk of death among all referred patients, the instrumental variable regression analyses estimated a substantially reduced risk of death for patients whose referrals were attributable to their physicians' referral threshold. Collectively, these findings suggest that raising the threshold for referral may substantially lower the use of specialised services, weakly reduce GP workload, and possibly have a decisive impact on patient safety through increased mortality for the affected patients.

Comparison with existing literature

To our knowledge, this is the first study assessing the effect of an acute referral to hospital from the out-of-hours services for patients with unclear indications for referral. Previous studies have found substantial differences between out-of-hours physician referral desicions, (24, 25) suggesting a varying threshold for referring. (26, 27) An essential aspect of variation in referral thresholds between physicians is for which patients this variation will apply. Most likely, a substantial share of the patients in primary care will never be directly affected by the physicians´ varying thresholds, as their referral needs are quite obvious. However, for some patients, the referral threshold of their treating physician will be decisive of their further care. It is also reasonable to believe that these patients are the ones most to be likely affected by general requests for primary care physicians to lower their referral rates. Increasing spendings on health services worldwide. (28, 29) (30) and challenges from crowded emergency departments (31) urge to reduce patient inflow to costly and limited specialised services. The incongruent views on referral practice within the health services serve as a base for conflict. (32, 33) Restricting referral opportunities and using GP referral rates as quality indicators have been debated. (34-36) GPs with high referral rates have been targeted to change practice to reduce pressure on specialised care, whereas there have

been less if any, such systematic efforts towards GPs with low referral rates to change their practice. Although reducing the pressure in the services is most likely positive and necessary, the impact on the patients whose entrance to specialised services is delayed or even denied by such changes in referral practice is challenging to study. Confounding by indication may put results from conventional observational studies in doubt. Our results suggest that patients with a referral attributable to their physicians' referral threshold are likely to affect health services with substantially higher hospital use and slightly higher use of outpatient specialist clinics in the time after referral. These patients are also more likely to use the GP services, contrary to what could be expected when the specialised services "handle the presented problem". Hence, such referrals do not relieve the GPs from workload.

Further, our results suggest a substantially reduced mortality, sustaining through the first six months after a referral attributable to physician's referral threshold. If the assumptions from our instrumental variable analyses hold, our results suggest that the potential impact from altering referral thresholds should be carefully assessed before implemented, justifying the use of randomized controlled trials. Our results also have implications for cases where the hospitals capacity limits are reached, announced as "red alert" or "black alert", where the threshold for referring is automatically raised. Our findings indicate that these situations may have detrimental effects on the affected patients. This is also in line with the findings of increased mortality for hip fracture patients who are discharged due to increased pressure/strain on the hospital. (37)

Strengths and limitations

We used an instrumental variable design in an attempt to address confounding by indication. Hence, by using the physician referral preference as an instrument to predict the patient's chance of referral, we could estimate the effects of referrals attributable to the physicians' referral threshold, and not to the patients' health conditions. To give valid results, our instrument should satisfy a set of assumptions. Although we had a large study, instrumental variable analyses are power sensitive, and the confidence intervals were wide. Our preference instrument showed weak or no associations with possible confounding variables, a result in favour of our independence assumption. There should be no effect of the physician referral preference and our outcomes other than via hospital referral for the

index patient. This assumption could be violated if referral preference was associated with specific treatment actions unrelated to the referral decision. However, as time is scarce in the out-of-hours setting, most treatment actions could more likely be seen as alternatives to referral.

We also investigated the effect of an alternative definition of our instrument. The results from these analyses provided similar results. We chose to include only elderly patients for our study, as we believe an appropriate referral decision to be both challenging and influential for these patients. Elderly patients are known to present with more diffuse symptoms, and be frailer regarding waiting time, transport, and changing environments. (38) If admitted to hospital, they are also prone to adverse events, (39) disabling, (40) and medication errors since they more often use multiple medications. (41) They are also more vulnerable to overdiagnostics, as they more often experience incidental findings in clinical, laboratory and radiological investigations, (42) where clinical guidelines may require follow up. Further, if admitted to the hospital, it may take time to arrange proper primary care to allow for discharge, demanding capacity. It is, therefore, most relevant to avoid unnecessary referrals for these patients.

Implications for future research

Referral thresholds have been thoroughly studied with the aim to reduce unnecessary hospital referrals and admission. The findings presented in this paper emphasise the need to consider the health services in the larger sense to avoid the pitfall of applying mending measures in one part without acknowledging the effects on the other parts of the service or the impact on patient safety. Our study support that referrals attributable to physicians' referral thresholds impose a high use of hospitals, possibly avoidable by raising referral thresholds, thus representing a target to reduce pressure on the specialised services. However, the finding of substantially reduced mortality for these patients, imply that at least for some of them, such referrals hold extensive value, and further that simply asking primary care physicians to raise their referral threshold could have detrimental consequences. Our findings prompt a need for further research to identify and evaluate more specific measures other than simply raising the thresholds for referral to reduce the increasing inflow of patients to specialised care, particularly considering the ageing and increasing population.

Improving opportunities to observe patients over time in out-of-hours services may improve the accuracy of referrals and prevent unnecessary referrals. Increasing the capacity in emergency departments for such observation could also be helpful, as prolonged observation is shown to lead to lower admission rates from the emergency departments. (43) However, our findings support the need for thorough evaluation of any measures taken to reduce referrals from primary care.

Conclusion

Out-of-hours patients whose referrals to hospital are attributable to their physicians' referral threshold contribute substantially to the use of health services. However, such referrals seem protective by reducing the risk of death during the first six months after the referral. This means that raising the thresholds for referral as a measure to lower pressure on overcrowded emergency departments and hospitals should not be encouraged without ensuring the quality and accuracy of the referral decisions. Our findings warrant further investigations, and the possible detrimental effects of changing thresholds for referral to hospital justifies thorough evaluation, such as randomised clinical trials.

Supplementary material Paper III

Supplementary material

I) Data extraction and organisation

II) Study design

III) Additional analyses & methods

I) Data extraction and organisation

This study is based on a large data set linking public healthcare registers in Norway to demographic information about the patients and characteristics of their primary care physicians. It comprises all contacts with the primary care and specialised services in Norway from 2008-2016, except contacts with, and admissions to psychiatric hospitals.

All patient contacts with Norwegian publicly funded general practitioner services generate a unique claim for reimbursement. These claims are submitted from the treating physician or service to the Norwegian Health Economics Administration and are registered in the Control and Payment of Health Reimbursement Register (KUHR). (14) For this study, we used the reimbursement claims data from all out-of-hours contacts in Norway from 2008-16, performed between 16:00 and 07:59 on weekdays, or during both day or night on Saturday, Sunday and public holidays. These claims include patient ID, time, patient diagnoses, a unique physician identification number, and the type of contact, (e.g., telephone, consultation, home-visits, including a specific code for out-of-hours work.) Claims missing the unique physician identification number were excluded. In some out-of-hours services, physicians have a fixed salary agreement, (mostly for working night-time, but some places also for late-shifts). In these cases, the claims are filed from the service or municipality, and will thus not include the unique physician ID. The share of claims filed in such ways, vary from year to year in our material. We further selected all claims containing the specific codes for out-of-hours work where physicians assessed patients, face-to-face or by telephone (codes: 2ak, 2fk, 1ak, 1bk, 1g). We excluded claims containing codes for home visits (codes: 2nk, 11ak, 11nk, 21k). We also excluded contacts where the patient's regular GP were present at the out-of-hours services, since such contacts could have been arranged between the patient and the GP. For this study we chose to study patients who were supposedly

unfamiliar to the out-of-hours physician and the staff. Therefore, we excluded claims where the patient met his or her regular GP, and where the patient had been assesses by the same physician earlier in the study period. Further, we excluded patient contacts from years where patients were frequent attenders, i.e., where they had more than four out-of-hours contacts per year (the 90-percentile for yearly contacts).

II) Study design

In this study our main aim was to estimate the effects of being referred to the hospital from out-of-hours services for patients where the indications for such referral were unclear. Further, we chose to concentrate on the effects for elderly patients. These are likely more challenging to assess and highly affected by a referral decision, as they often are frailer. To increase the contrast to the regular GP setting, we chose to only include patients who are presumably unknown to the physician and staff, as described above. We intended to investigate the effects on both patient outcomes and the dynamics of health services. Thus, our outcomes of interest were defined as health service use and mortality following the referral, both short-term and long-term. Based on this, we defined the follow-up period to be 0-10, 0-30, 0-90 and 180 days.

This study design is based on the main assumption that there are no systematic associations between the treating physician's referral preference and the potentially confounding patient factors in each index contact. By including only elderly patients, assessed by the particular out-of-hours physician for the first time, and excluding contacts from years where patients had >4 contacts, the patients were more similar, thus decreasing the risk of such systematic differences. An example of an association we are trying to avoid, would be that a physician with high referral preference would be more likely to meet patients with more severe conditions. Out-of-hours services provide a good study setting, since patients usually do not choose when to get acutely ill, and we can assume that in most cases, they have no knowledge of which physicians are on-call in their area. There is often only one physician serving an area per time in the Norwegian out-of-hours services; however, there is a trend towards centralizing the out-of-hours services, resulting in larger out-of-hours services, with several physicians on-call at the same time.

Even if systematic differences in patient and physician characteristics are less likely in the out-of-hours setting than in the regular GP setting, there are still some possible confounding factors that need to be addressed:

1) **Local differences between the out-of-hours services, adaptation to local conditions**

We assume that there are substantial differences between the various out-of-hours services across Norway. These are most likely based on local adaptations to the patient population's needs, and to the other parts of the services, such as distance to hospital, and transportation resources. Hence, the provision of out-of-hours service in a city differs from the out-of-hours services in a rural municipality. Based on this knowledge, we only compare contacts within the same services. In Norway, all municipalities have a unique municipality code. The extensive collaboration between municipalities leads to shifting geographical locations of many out-of-hours services, especially in the scarcely populated municipalities. The KUHR claims lack information about the actual geographical location where the contact took place. Although some patients fall ill when travelling, we assume that patients visit the out-of-hours services in their home area in most of the cases. However, the physicians will mainly work within only one out-of-hours service at the time, although the same out-of-hours service may often serve the population from several municipalities. We handled this problem by linking the unique patient ID from each claim to the municipality code where patients were registered as residents (Statistics Norway). To define the location of the out-of-hours services, we used the municipality codes´ modal value among patients within physicians per week. As a result, the municipality where most of the patients (within one week, seeing the same physician) live was defined as the municipality where the contacts for this physician took place that particular week. Thus, this could change through the time periods in the data material. If the value was the same for two or more municipalities, we chose the lower code, as this often represents a larger municipality.

2) **Possible time trends over the years**

To handle possible changes in both organisational factors and referral practices over time, we only compare patients within the same year.

3) **Systematic differences between day shifts, late shifts and night shifts**

Based on the assumption that there can be systematic differences between daytime, afternoon and nighttime in patients contacts and physicians staffing the out-of-hours-services, we chose to divide 24 hours into three time units; late shift (16:00-23:59); night shift (00:00-07:59); day shifts (08:00-15:59) during weekends or holidays. Thus, we only compare patients contacting the out-of-hours service in the same out-of-hours shift during the day.

Since our instrument was based on the patients' sex, we compared only female patients with other female patients, and male patients with other male patients.

Based on the assumptions above, we matched patients in groups defined by combining information on patients of the same sex visiting the same out-of-hours station in the same out-of-hours shifts within the same year. For example, we compared all female patients visiting the same out-of-hours station all late shifts (16:00-23:59) in 2015. By analysing only within-group variability, we effectively controlled for all confounding that was constant within each group. To avoid the effect of possible patient selection in situations where two or more GPs were on-call at the same time, we used the weighted average of GP characteristics within each out-of-hours shift in each service.

III) Additional analyses

Table S1 Balance test of confounders

Table S1 Balance test of confounders. Regression analysis showing the association between potential confounders (patients characteristics) and the physician´s preference for referral to hospital.

n = 922,796	Instrumental variable analyses* (unscaled)			Instrumental variable analyses* (scaled)**			Multivariable adjusted analyses***		
	Beta	**LCI**	**UCI**	**Beta**	**LCI**	**UCI**	**Beta**	**LCI**	**UCI**
Age in years	**0.20**	-0.23	0.63	**0.61**	-0.32	1.54	**1.48**	1.44	1.52
Low education	**-0.01**	-0.02	0.00	**-0.02**	-0.04	0.01	**0.01**	0.00	0.01
Immigration status	**0.00**	0.00	0.01	**0.01**	0.00	0.02	**0.00**	0.00	0.00
ICD-10 Chapter III (Circulatory system)	**0.00**	-0.01	0.01	**-0.01**	-0.03	0.01	**0.01**	0.01	0.01
ICD-10 Chapter IX (Neoplasms)	**0.00**	-0.01	0.00	**0.02**	-0.01	0.05	**0.01**	0.01	0.01
Health care use previous month:									
Elective hospital admission	**0.00**	-0.01	0.01	**0.01**	-0.02	0.03	**0.02**	0.02	0.02
Unplanned hospital admission	**0.00**	-0.01	0.02	**0.00**	-0.02	0.01	**0.03**	0.03	0.03
General practitioner visit	**0.00**	-0.01	0.01	**0.00**	-0.02	0.02	**0.00**	0.00	0.00
Outpatient clinic visit	**0.00**	-0.01	0.00	**0.00**	-0.02	0.01	**0.02**	0.02	0.02

***Adjusted for patient age, and age^2, month and weekday.**

****Scaled according to the strength of the instrument-exposure association.**

*****Adjusted for patient age, and age^2, year, month and weekday and hour.**

Balance tests/confounder analyses

To justify our assumption of no association between the physician referral preference and the patient characteristics for the index contacts within our defined groups, we performed balance tests where we calculated the associations between selected patient characteristics known as potential confounders, and the physician referral preference. These are shown in Supplementary Table 1. We made variables for the healthcare use prior to the out-of-hours contact for all patients. We also made variables based on being discharged with specific diagnoses; 1) ICD-10 Chapter IX, Diseases of the circulatory system and 2) ICD-10 Chapter III, Neoplasms (Malignant diseases). Together with variables for patient sex, age, education level and immigration status, these variables were used in analyses to test whether patient characteristics affected which GP they saw when visiting the out-of-hours services. As we can see from the results from Supplementary Table 1, there was little or no evidence of any differences between GP characteristics and possible patient confounders, given our design. Hence, these results provide a strong support of the confounder independence assumption of our design.

Table S2 Instrumental variable estimates for difference in health services use using physician referral preference as an instrumental variable

Table S2 Instrumental variable analyses estimates of differences in days with health services use for patients referred (vs not referred) to hospital within 10 hours of an out-of-hours contact*					
Primary care physician	**Mean difference**	**LCI**	**UCI**	**F-statistics**	**P-value endogeneity test****
Within 10 days	**-0.13**	-0.2	-0.06	1,212	<0.001
Within 30 days	**0.05**	-0.08	0.19	1,212	<0.001
Within 90 days	**0.44**	0.15	0.74	1,215	0.037
Within 180 days	**0.46**	0.17	0.76	1,199	0.037
Outpatient clinics					
Within 10 days	**0.33**	0.29	0.37	1,212	0.013
Within 30 days	**0.39**	0.32	0.47	1,212	0.111
Within 90 days	**0.64**	0.49	0.8	1,215	0.591
Within 180 days	**0.78**	0.53	1.04	1,199	0.335
Days in hospital					
Within 10 days	**3.25**	3.08	3.42	1,212	<0.001
Within 30 days	**3.58**	3.27	3.89	1,212	<0.001
Within 90 days	**3.81**	3.28	4.34	1,215	<0.001
Within 180 days	**3.88**	3.18	4.58	1,199	<0.001

* Instrument *'physician referral preference'* is defined as the share of the physicians' out-of-hours contacts who were referred in the period, for the opposite sex. F-statistics estimates the strength of the instrument. Adjusted for patient age, age^2, month, and weekday

** Test of difference in estimates between multivariable adjusted and instrumental variable analyses.

Table S3 Instrumental variable estimates for an alternative definition of the instrument ('physicians referral preference 2')

Table S3 Instrumental variable analyses estimates of differences in days with health services use for patients referred (vs not referred) to hospital within 10 hours of an out-of-hours contact, using an alternative definition of the instrument*

Primary care physician	Mean difference	LCI	UCI	F-statistics
Within 10 days	**-0.20**	-0.30	-0.09	1,212
Within 30 days	**-0.06**	-0.26	0.14	1,212
Within 90 days	**0.30**	-0.12	0.73	1,215
Within 180 days	**0.23**	-0.19	0.66	1,199
Outpatient clinics				
Within 10 days	**0.37**	0.31	0.43	1,212
Within 30 days	**0.45**	0.34	0.55	1,212
Within 90 days	**0.63**	0.39	0.86	1,215
Within 180 days	**0.88**	0.48	1.28	1,199
Days in hospital				
Within 10 days	**3.30**	3.06	3.54	1,212
Within 30 days	**3.70**	3.25	4.14	1,212
Within 90 days	**3.98**	3.24	4.73	1,215
Within 180 days	**4.01**	3.02	4.99	1,199

*Instrument *'physician referral preference 2'* is defined as the share of the physicians' out-of-hours contacts who were referred in the period, but excluding the year of the index contact, and one year before and after. F-statistics estimates the strength of the instrument. Adjusted for patient age, age^2, month, and weekday.

1. NHS. Online version of the NHS long term plan 2020 [Available from: https://www.longtermplan.nhs.uk/online-version/chapter-1-a-new-service-model-for-the-21st-century/2-the-nhs-will-reduce-pressure-on-emergency-hospital-services/.
2. Lasserson D, Smith H, Garland S, Hunt H, Hayward G. Variation in referral rates to emergency departments and inpatient services from a GP out of hours service and the potential impact of alternative staffing models. Emerg Med J. 2021.
3. Velasco Garrido M, Zentner A, Busse R. The effects of gatekeeping: a systematic review of the literature. Scand J Prim Health Care. 2011;29(1):28-38.
4. Sripa P, Hayhoe B, Garg P, Majeed A, Greenfield G. Impact of GP gatekeeping on quality of care, and health outcomes, use, and expenditure: a systematic review. The British journal of general practice : the journal of the Royal College of General Practitioners. 2019;69(682):e294-e303.
5. Roland M, Abel G. Reducing emergency admissions: are we on the right track? BMJ. 2012;345:e6017.
6. Hernan MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? Epidemiology. 2006;17(4):360-72.
7. Boef AG, le Cessie S, Dekkers OM, Frey P, Kearney PM, Kerse N, et al. Physician's Prescribing Preference as an Instrumental Variable: Exploring Assumptions Using Survey Data. Epidemiology. 2016;27(2):276-83.
8. Schneeweiss S, Setoguchi S, Brookhart A, Dormuth C, Wang PS. Risk of death associated with the use of conventional versus atypical antipsychotic drugs among elderly patients. CMAJ. 2007;176(5):627-32.
9. Wang PS, Schneeweiss S, Avorn J, Fischer MA, Mogun H, Solomon DH, et al. Risk of death in elderly users of conventional vs. atypical antipsychotic medications. N Engl J Med. 2005;353(22):2335-41.
10. Svedahl ER, Pape K, Austad B, Vie GA, Anthun KS, Carlsen F, et al. Effects of GP characteristics on unplanned hospital admissions and patient safety. A 9-year follow-up of all Norwegian out-of-hours contacts. Fam Pract. 2021.
11. Saunes IS, Karanikolos M, Sagan A. Norway: Health system review. Health systems in Transition, 2020. World Health Organization. Regional Office for Europe, European Observatory on Health Systems and Policies 2020.
12. Berchet C, Nader C. The organisation of out-of-hours primary care in OECD countries. Paris; 2016. Report No.: 89.
13. Nieber T HE, Bondevik GT, Hunskår S, Blinkenberg J, Thesen J, Zakariassen E. . Organisering av legevakt [Organization of Norwegian out-of-hours primary health care services]. Norwegian. . Tidsskr Nor Laegeforen 2007 (May 17;127(10):1335-8.).
14. The Norwegian Directorate of Health. Control and Payment of Health Reimbursement Register Data base 2019 (KUHR-databasen 2019) 2019 [Available from: https://www.helsedirektoratet.no/tema/statistikk-registre-og-rapporter/helsedata-og-helseregistre/kuhr.
15. Bakken IJ, Ariansen AMS, Knudsen GP, Johansen KI, Vollset SE. The Norwegian Patient Registry and the Norwegian Registry for Primary Health Care: Research potential of two nationwide health-care registries. Scand J Public Health. 2020;48(1):49-55.
16. Statistics Norway. Statistisk sentralbyrå 2021 [Available from: https://www.ssb.no/en.

17. Norwegian Insitute of Public Health. The Norwegian Cause of Death Registry 2020 [Available from: https://www.fhi.no/en/hn/health-registries/cause-of-death-registry/.
18. The Norwegian Directorate of eHealth. The regular general practitioner register (Fastlegeregisteret) 2020 [Available from: https://helsedata.no/forvaltere/helsedirektoratet/fastlegeregisteret/.
19. Sandvik H, Hunskaar S. Frequent attenders at primary care out-of-hours services: a registry-based observational study in Norway. BMC Health Serv Res. 2018;18(1):492.
20. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. Stat Methods Med Res. 2017;26(5):2333-55.
21. Swanson SA, Hernan MA. Commentary: how to report instrumental variable analyses (suggestions welcome). Epidemiology. 2013;24(3):370-4.
22. Correia S. A Feasible Estimator for Linear Models with Multi-Way Fixed Effects* 2016 [Available from: http://scorreia.com/research/hdfe.pdf.
23. Inoue A, Solon G. Two-Sample Instrumental Variables Estimators. Review of Economics and Statistics. 2010;92(3):557-61.
24. Rossdale M, Kemple T, Payne S, Calnan M, Greenwood R. An observational study of variation in GPs' out-of-hours emergency referrals. British Journal of General Practice. 2007;57(535):152-4.
25. O'Donnell CA. Variation in GP referral rates: what can we learn from the literature? Family Practice. 2000;17(6):462-71.
26. Calnan M, Payne S, Kemple T, Rossdale M, Ingram J. A qualitative study exploring variations in GPs' out-of-hours referrals to hospital. The British journal of general practice : the journal of the Royal College of General Practitioners. 2007;57(542):706-13.
27. Ingram JC, Calnan MW, Greenwood RJ, Kemple T, Payne S, Rossdale M. Risk taking in general practice: GP out-of-hours referrals to hospital. The British journal of general practice : the journal of the Royal College of General Practitioners. 2009;59(558):e16-24.
28. National Health Service. A&E attendances and emergency admissions. 2021 [Available from: https://www.england.nhs.uk/statistics/statistical-work-areas/ae-waiting-times-and-activity/.
29. Eric C. Schneider AS, Michelle M. Doty, Roosa Tikkanen, Katharine Fields, Reginald D. Williams II. Mirror, Mirror 2021: Reflecting Poorly: The Commonwealth Fund; 2021 [Available from: https://www.commonwealthfund.org/publications/fund-reports/2021/aug/mirror-mirror-2021-reflecting-poorly.
30. OECD. Health at a glance 2021 [Available from: https://www.oecd.org/health/health-at-a-glance/.
31. Morris ZS, Boyle A, Beniuk K, Robinson S. Emergency department crowding: towards an agenda for evidence-based intervention. Emerg Med J. 2012;29(6):460-6.
32. Oliver D. David Oliver: How hospital doctors' own pressures affect GPs. BMJ. 2021;374:n2358.
33. Office of the Auditor General of Norway (Riksrevisjonen). Riksrevisjonens undersøkelse av myndighetenes arbeid med å sikre god henvisningspraksis fra fastlegene til spesialisthelsetjenesten 2018.
34. Gulland A. All GP referrals should be subject to clinical peer review, says NHS England. BMJ. 2017;358:j4106.
35. Hawkes N. BMA meeting: GPs are being told to limit their hospital referrals, meeting hears. BMJ. 2011;342:d4093.

36. Love T, Dowell AC, Salmond C, Crampton P. Quality indicators and variation in primary care: modelling GP referral patterns. Fam Pract. 2004;21(2):160-5.
37. Nilsen SM, Bjorngaard JH, Carlsen F, Anthun KS, Johnsen LG, Vatten LJ, et al. Hospitals Discharge Tendency and Risk of Death - An Analysis of 60,000 Norwegian Hip Fracture Patients. Clin Epidemiol. 2020;12:173-82.
38. Linn Solveig Sortland LMHoIS. Eldre i den akuttmedisinske kjeden
Rapport nr. 1-2021 Nasjonalt kompetansesenter for legevaktmedisin. NORCE; 2021.
39. Forster AJ, Asmis TR, Clark HD, Al Saied G, Code CC, Caughey SC, et al. Ottawa Hospital Patient Safety Study: incidence and timing of adverse events in patients admitted to a Canadian teaching hospital. CMAJ. 2004;170(8):1235-40.
40. Sager MA. Functional Outcomes of Acute Medical Illness and Hospitalization in Older Persons. Archives of Internal Medicine. 1996;156(6).
41. Perez T, Moriarty F, Wallace E, McDowell R, Redmond P, Fahey T. Prevalence of potentially inappropriate prescribing in older people in primary care and its association with hospital admission: longitudinal study. BMJ. 2018;363:k4524.
42. Skinner TR, Scott IA, Martin JH. Diagnostic errors in older patients: a systematic review of incidence and potential causes in seven prevalent diseases. Int J Gen Med. 2016;9:137-46.
43. Asheim A, Nilsen SM, Carlsen F, Eide Naess-Pleym L, Uleberg O, Dale J, et al. The effect of emergency department delays on 30-day mortality in Central Norway. Eur J Emerg Med. 2019;26(6):446-52.

Appendix I

Hovedspørsmål/tema	Supplerende/utdypende ved behov
«På et legekontor er det ofte travelt, og dere har helt sikkert alle opplevd arbeidspress på jobben. I dette gruppeintervjuet er arbeidspress hovedtema, og vi er interessert i å finne ut hvordan dere opplever dette i deres arbeids hverdag. Vi er ute etter legenes/sekretærenes erfaringer og refleksjoner i tilknytning til både arbeidspress og tidspress. Vi antar at dette innebærer mange ulike mange ulike synspunkter og oppfatninger. Før vi begynner vil vi understreke at det selvfølgelig ikke fins noen riktige og gale svar; alle innspillene fra dere er interessante for oss. Det er også interessant å få fram hva dere er enige i og hvor oppfatningene spriker». ***«Har dere noen spørsmål før vi begynner?»***	
Hvordan vil dere beskrive legekontoret?	Sammenlignet med andre? Travelt?
«Nå kommer noen spørsmål som omhandler deres egen praksis/legekontor»	
Hvordan er situasjonen nå for tiden?	I din egen praksis/liste, på legesenteret
Kan du beskrive en normal/vanlig/typisk dag vs. en spesielt travel dag på jobb?	
Hva er de vanligste situasjonene som gjør at du føler deg presset på tid?	Der og da, i løpet av en arbeidsdag Over lengre tid Hva skal til for at stort tidspress oppleves som et problem Foruten influensasesongen – er det andre perioder som er travlere enn andre
Hvilke konsekvenser har tidspress for deg/dere?	For deg? For kollegaene dine? For pasientene? Spiser du vanligvis lunsj? Rekker dere å snakke sammen? På kort sikt (samme dag), på lengre sikt Motivasjon? Mestringsfølelse? Følelsen av å gi pasienten god behandling/oppfølging?
Hva slags støtte får du av de andre på legekontoret når du er presset på tid?	I hvilken grad finner man støtte og motivasjon hos kolleger på tvers av faggrupper (altså: hvor viktig er fastlegene for spl/sekretær og omvendt?
«Tenk deg nå en dag med spesielt stort tidspress (gjenta situasjoner som er nevnt tidligere, gi eksempler)» eks forsinkelser pga pasient som ble akutt dårlig/pasient som måtte innlegges akutt i psykiatrien/pasient som trengte akutt sykebesøk, syk kollega, siste dag før ferie…	
Hvordan håndterer du en slik situasjon? **Hvilke prioriteringer gjøres?** (Viktig her å registrere i hvilken retning. Ex: ja, økt tidspress påvirker sannsynligheten for å henvise videre, men øker eller reduseres	(Leger) I hvilken grad får dette konsekvenser for: – Hvorvidt du velger å henvise en pasient til spesialist eller ikke – Hvorvidt du legger en pasient inn på sykehus eller ikke – Hvorvidt du gjennomfører sykebesøk – Hvorvidt du får besvart henvendelser fra hjemmesykepleien

sannsynligheten? (den øker nok..) Samme med kontrolltime. Setter opp flere når travelt. Som igjen øker ventetid og arbeidspress på sikt)	– Hvorvidt du prioriterer dialogmøter/ansvarsgruppemøter – Omfanget av behandling som blir igangsatt der og da – Ta tid til å samtale med folk i krisesituasjoner – Igangsetting av kognitiv terapi, støttesamtaler ved lettere psykiske lidelser – Hvorvidt du avtaler/setter opp kontrolltime – Forskrivning av B-preparater (tar ofte lengre tid å si nei enn ja) – Sykemeldinger (der man kanskje kunne funnet andre løsninger) – Fornyelse av resepter elektronisk der du ideelt sett hadde ønsket å ha folk på time – Utskrivning av attester uten å ha folk på time (Sekretærer) I hvilken grad får dette konsekvenser for: – Hvorvidt pasienter som ringer inn får tilbud om time, evt legges som memo til legen – Ivaretakelsen av pasientene Går ut over spesielle pasientgrupper (diagnoser, andre kjennetegn)?
«Tenk deg nå en periode med spesielt stort tidspress (gjenta situasjoner som er nevnt tidligere, gi eksempler)» eks en av legene /kollegaene har brukket et ben og vært borte fra jobben i tre uker uten at dere har fått inn vikar, influensasesong, ny medarbeider …	
Hvordan håndterer du en slik situasjon? **Hvilke prioriteringer gjøres?**	Samme underspørsmål som over
«I fortsettelsen av prosjektet vårt ønsker vi å identifisere situasjoner som gir tidspress, slik at vi deretter kan benytte disse sitasjonene for å se på konsekvenser av tidspress. Vi har derfor et par spørsmål om konkrete situasjoner som vi er interessert i å vite mer om»	
Hvilke pasientgrupper er det som skaper mest tidspress?	Hva er årsaken til at akkurat disse skaper press? Spør om psykiske lidelser og vansker spesielt + asylsøkere (pasienter som prater mye? Omstendelige – tar opp flere problemer på samme time? Hører dårlig? Kravstore - trenger hjelp NÅ uansett problemstilling) For- og etterarbeid; hvilken type arbeid størst belastning, minst motiverende
Sykdom hos fastlegen, kurs, vikarer - i hvilken grad skaper slike situasjoner tidspress?	NB: Hvordan påvirkes arbeidspresset av at man har vært borte fra praksis en periode (kurs, ferie, sykdom)?
«Nå har vi snakket mye om det som gir tidspress og arbeidspress i arbeidsdagene deres. Tidligere undersøkelser fra allmennpraksis har likevel funnet at leger stort sett er tilfredse med jobbsituasjonen sin og også klarer å håndtere stort trykk over tid på en god måte. Tenk derfor igjen på situasjonen sånn alt-i-alt»	
Slik situasjonen er nå; hvor stor utfordring er tidspress/arbeidspress for deg?	Endring over tid? Påvirker det evt. motivasjonen for å fortsette i jobben/ i allmennpraksis ? Risiko for å gjøre feil?

Hva tenker dere kan bidra til å redusere tidspresset i fastlegetjenesten? **Hva gjør dere for å spare tid?**	Lønns- og arbeidsforhold mv. Kutte pasientlister? Tydeligere beskjeder til pasienter? Mulighetene til å henvise til kommunehelsetjenesten vs. spesialisthelsetjenesten Man skal nå i gang med forsøk med primærhelseteam – hva tenker dere om det?
«Et av målene med prosjektet vårt er å se på hvordan press i primærhelsetjenesten og spesialisthelsetjenesten påvirker hverandre. Til slutt har vi derfor noen spørsmål om deres erfaringer i forhold til det som skjer når pasienter blir utskrevet fra sykehuset, og særlig eldre pasienter med behov for oppfølging».	
Hva er din/deres generelle oppfatning av samhandling med sykehuset i forhold til pasienter som blir utskrevet? **Har du opplevd at (eldre) pasienter har blitt utskrevet for tidlig fra sykehuset? - beskriv i så fall hva som skjedde? Når skjedde det sist?**	Hvordan får du vanligvis informasjon om at pasienten er kommet hjem? Hvordan får du informasjon om behovet for oppfølging etter en innleggelse? Hvilke rutiner har du for eldre pasienter som skrives ut fra sykehuset? Er det noen forhold som kunne vært forbedret i forhold til samhandling med sykehuset i forbindelse med utskrivelse?

For hver deltaker noteres alder, kjønn, spesialisering og forhold ved legekontoret (antall ansatte i ulike stillinger, oppgaver delegert til personell, rutiner for timebestilling (internett/sms/tlf, ØH-timer, time samme dag), pasientsammensetning og håndtering av elektronisk kommunikasjon; mengde epikriser/PLO-meldinger/SMS fra pasienter som skal følges opp.

Intervjuguide - hovedspørsmål

arbeidspress hovedtema, finne ut hvordan dere opplever dette i deres arbeids hverdag.

Hvordan vil dere beskrive legekontoret? (sammenlignet med andre)

egen praksis

Hvordan er situasjonen nå for tiden?

Kan du beskrive en normal/vanlig/typisk dag vs. en spesielt travel dag på jobb?

Hva er de vanligste situasjonene som gjør at du føler deg presset på tid? (dag, periode, hvorfor problem)

Hvilke konsekvenser har tidspress for deg/dere? (selv, kolleger, pasienter, kort og lang sikt, motivasjon)

Hva slags støtte får du av de andre på legekontoret når du er presset på tid?

eksempel på situasjon - dag

Hvordan håndterer du en slik situasjon?

Hvilke prioriteringer gjøres?

eksempel på situasjon - periode

Hvordan håndterer du en slik situasjon?

Hvilke prioriteringer gjøres?

Henvisning
Innleggelse
Sykebesøk
Hjemmesykepleien
Møter
Omfang av behandling
Samtalebehandling
Avtale kontroll
Foreskrivning medisiner (b-prep)
Sykmelding
Resepter/attester uten time
Går det ut over spesielle pasientgrupper?

konkrete situasjoner

Hvilke pasientgrupper er det som skaper mest tidspress?

Sykdom hos fastlegen, kurs, vikarer - i hvilken grad skaper slike situasjoner tidspress?

alt-i-alt

Slik situasjonen er nå; hvor stor utfordring er tidspress/arbeidspress for deg? (over tid, motivasjon, gjøre feil)

Hva tenker dere kan bidra til å redusere tidspresset i fastlegetjenesten? (lønn, liste, pasienter, andre aktører)

samhandling

Hva er din/deres generelle oppfatning av samhandling med sykehuset i forhold til pasienter som blir utskrevet? (hvordan skjer dette vanligvis, muligheter for forbedring)

Har du opplevd at (eldre) pasienter har blitt utskrevet for tidlig fra sykehuset? - beskriv i så fall hva som skjedde? Når skjedde det sist?

Appendix II

Forespørsel om deltakelse i forskningsprosjekt

Opplevelse av arbeidspress i fastlegepraksis – gruppeintervjuer med fastleger og helsesekretærer.

Bakgrunn og formål
Denne kvalitative delstudien inngår i forskningsprosjektet *«Health care services under pressure – consequences for patient flows, efficiency and patient safety»*, finansiert av Norges Forskningsråd. Prosjektet utføres av forskere ved St. Olavs Hospital og NTNU og gjennomføres i perioden 1. januar 2017 til 31. desember 2020 (https://www.researchgate.net/project/Health-care-services-under-pressure-Consequences-for-patient-flows-efficiency-and-patient-safety-in-Norway)

Fastlegeordningen legger stort ansvar på den enkelte fastlege for å yte adekvate tjenester for sin listepopulasjon. Over tid synes det å ha skjedd en gradvis utvidelse av fastlegenes ansvar og arbeidsoppgaver. Dette er en utvikling vi kan vente vil fortsette, sett i lys av samhandlingsreformen og en stadig eldre befolkning. Arbeidsoppgaver- og mengde vil derfor øke for allmennlegetjenesten i årene som kommer. Vi vet imidlertid lite om hvordan norske fastleger opplever arbeidspress, og hvilke konsekvenser dette har – for legene selv, for deres praksis og for deres pasienter/listepopulasjon.

I dette prosjektet ønsker vi derfor å innhente informasjon direkte fra praksisfeltet gjennom gruppeintervjuer med fastleger og helsesekretærer. Formålet med intervjuene er å undersøke hvordan dere opplever arbeidspress i deres praksis, og hvilke tanker dere har omkring mulige konsekvenser av slikt press. Den innsamlede informasjonen fra intervjuene vil bli analysert i etterkant. Resultatene vil gi oss verdifull og ny kunnskap, og målet er å publisere disse i en vitenskapelig artikkel. Resultatene vil også benyttes til å justere eller generere nye forskningshypoteser for resten av det store prosjektet vårt, hvor vi i hovedsak vil benytte store registerbaserte datamaterialer. Informasjon om hvilke situasjoner som gir økt arbeidspress vil være spesielt interessant, ettersom et sentralt mål med de kvantitative undersøkelsene er å finne gode indikatorer på arbeidspress. I tillegg håper vi å kunne bidra til kunnskapsutveksling mellom klinikk og forskning ved å holde dere som studiedeltakere orientert om de resultatene som prosjektet frembringer. Alle prosjektdeltakere som ønsker det vil motta artikler eller rapporter som blir publisert.

Hva innebærer deltakelse i studien?
Deltakelse i studien innebærer at du deltar på et gruppeintervju med flere andre fastleger/helsesekretærer. Intervjuet vil vare i ca. 60 minutter. Spørsmålene vil i hovedsak omhandle dine erfaringer og opplevelser med arbeidspress i din praksis, og krever ingen forberedelse.

Hvem er vi?
Prosjektgruppen vår er tverrfaglig sammensatt, og flere leger er med. I tillegg har vi et samarbeid med fastleger som er ansatt ved Institutt for samfunnsmedisin og sykepleie, NTNU. På intervjuet vil du møte Kristine Pape, Silje Kaspersen, Lena Janita Skarshaug, Ellen Rabben Svedahl og Marlen Toch-Marquardt som alle er forskere i prosjektet.

Hva skjer med informasjonen om deg?
Intervjuene vil tas opp på lydbånd og transkriberes. Du trenger ikke oppgi navn, hvilken kommune du bor i eller fastlegekontor i intervjuet. Datamaterialet lagres på sikre servere der kun prosjektdeltakerne har tilgang. Vi vil også registrere noen personopplysninger slik som kjønn, alder (ti-års kategorier), antall år som fastlege/helsesekretær (i kategorier) og spesialisering (for legene). I tillegg vil vi innhente noen opplysninger om legekontoret du jobber på (antall ansatte i ulike stillinger, oppgaver delegert til personell, rutiner for timebestilling, pasientsammensetning og håndtering av elektronisk kommunikasjon). Alle personopplysninger vil bli behandlet konfidensielt og kun prosjektgruppen vil ha tilgang til disse. Ditt navn, epostadresse eller telefonnummer vil lagres atskilt fra lydopptakene, og slettes ved prosjektslutt. Datainnsamlingen gjennomføres i 2017.

Studiens resultater vil bli publisert i flere forskningsrapporter. Rapportene vil ikke inneholde opplysninger som kan føres tilbake til deg eller din praksis. Prosjektet skal etter planen avsluttes **31.12.20**. Alle opptak og eventuelle personopplysninger vil da bli slettet.

Frivillig deltakelse
Det er frivillig å delta i studien, og du kan når som helst trekke ditt samtykke uten å oppgi noen grunn. Dersom du trekker deg, vil alle opplysninger om deg bli anonymisert, på lik linje med om du deltar.

Dersom du ønsker å delta eller har spørsmål til studien, ta kontakt med Kristine Pape – kristine.pape@ntnu.no eller 95117920

Forskningsprosjektet er tilrådd av Personvernombudet for forskning, referansenummer 54945

Samtykke til deltakelse i studien

På intervjutidspunktet vil vi samle inn skriftlig samtykke til deltakelse i forskningsprosjektet.

Jeg har mottatt informasjon om studien, og er villig til å delta

---
(Signert av prosjektdeltaker, dato)

Supplerende opplysninger om deltakere:
Følgende papirskjema fylles ut (utkrysning, noe tekst) av deltakerne selv i forbindelse med gruppeintervjuet. Opplysningene fra papirskjemaet vil sammenholdes og grovkategoriseres før de lagres elektronisk under et deltakernummer. Papirskjemaene vi deretter makuleres. Opplysningene vil via deltakernummeret kunne knyttes til stemmene på lydbånd, men kategoriseres slik at muligheten for å identifisere enkeltpersoner vil være liten. Dette deltakernummeret vil ikke kunne knyttes til kontaktinformasjon for deltakerne. Det vil ikke bli mulig å identifisere den enkelte lege når data blir publisert.

Legekontor		id

Kjønn
- ☐ Kvinne
- ☐ Mann

Alder
- ☐ 20-29 år
- ☐ 30-39 år
- ☐ 40-49 år
- ☐ 50-59 år
- ☐ 60-69 år (og eldre)

Antall år som fastlege
- ☐ Under 2 år
- ☐ 2-4 år
- ☐ 5-9 år
- ☐ 10-19 år
- ☐ 20 år eller mer

Spesialist
- ☐ Nei
- ☐ Ja, i allmennmedisin
- ☐ Ja, annen spesialitet

Antall pasienter på lista
- ☐ < 900
- ☐ 900-999
- ☐ 1000-1099
- ☐ 1100-1199
- ☐ 1200-1299
- ☐ 1300-1399
- ☐ 1400-1499
- ☐ 1500 eller mer

Antall dager pr. uke på legekontoret	
Antall pasienter pr. dag (ca)	
Oppsatt tid pr. pasient (ordinær time, ø-hjelp)	

Spesielle oppgaver (angi timer i uka hvis aktuelt)

☐ Kommunale oppgaver	Hva?	
☐ Deltar i legevakt	t/uke	
☐ Tillitsverv i legeforeningen el	Hva?	
☐ Spesielt ansvar i drift av legekontoret (f.eks IT-ansvar)	Hva?	
☐ Annet	Hva?	

Forespørsel om deltakelse i forskningsprosjekt

Opplevelse av arbeidspress i fastlegepraksis – gruppeintervjuer med fastleger og helsesekretærer.

Bakgrunn og formål
Denne kvalitative delstudien inngår i forskningsprosjektet *«Health care services under pressure – consequences for patient flows, efficiency and patient safety»*, finansiert av Norges Forskningsråd. Prosjektet utføres av forskere ved St. Olavs Hospital og NTNU og gjennomføres i perioden 1. januar 2017 til 31. desember 2020 (https://www.researchgate.net/project/Health-care-services-under-pressure-Consequences-for-patient-flows-efficiency-and-patient-safety-in-Norway)

Fastlegeordningen legger stort ansvar på den enkelte fastlege for å yte adekvate tjenester for sin listepopulasjon. Over tid synes det å ha skjedd en gradvis utvidelse av fastlegenes ansvar og arbeidsoppgaver. Dette er en utvikling vi kan vente vil fortsette, sett i lys av samhandlingsreformen og en stadig eldre befolkning. Arbeidsoppgaver- og mengde vil derfor øke for allmennlegetjenesten i årene som kommer. Vi vet imidlertid lite om hvordan norske fastleger opplever arbeidspress, og hvilke konsekvenser dette har – for legene selv, for deres praksis og for deres pasienter/listepopulasjon.

I dette prosjektet ønsker vi derfor å innhente informasjon direkte fra praksisfeltet gjennom gruppeintervjuer med fastleger og helsesekretærer. Formålet med intervjuene er å undersøke hvordan dere opplever arbeidspress i deres praksis, og hvilke tanker dere har omkring mulige konsekvenser av slikt press. Den innsamlede informasjonen fra intervjuene vil bli analysert i etterkant. Resultatene vil gi oss verdifull og ny kunnskap, og målet er å publisere disse i en vitenskapelig artikkel. Resultatene vil også benyttes til å justere eller generere nye forskningshypoteser for resten av det store prosjektet vårt, hvor vi i hovedsak vil benytte store registerbaserte datamaterialer. Informasjon om hvilke situasjoner som gir økt arbeidspress vil være spesielt interessant, ettersom et sentralt mål med de kvantitative undersøkelsene er å finne gode indikatorer på arbeidspress. I tillegg håper vi å kunne bidra til kunnskapsutveksling mellom klinikk og forskning ved å holde dere som studiedeltakere orientert om de resultatene som prosjektet frembringer. Alle prosjektdeltakere som ønsker det vil motta artikler eller rapporter som blir publisert.

Hva innebærer deltakelse i studien?
Deltakelse i studien innebærer at du deltar på et gruppeintervju med flere andre fastleger/helsesekretærer. Intervjuet vil vare i ca. 60 minutter. Spørsmålene vil i hovedsak omhandle dine erfaringer og opplevelser med arbeidspress i din praksis, og krever ingen forberedelse.

Hvem er vi?
Prosjektgruppen vår er tverrfaglig sammensatt, og flere leger er med. I tillegg har vi et samarbeid med fastleger som er ansatt ved Institutt for samfunnsmedisin og sykepleie, NTNU. På intervjuet vil du møte to-tre personer fra prosjektgruppen, som består av forskerne Kristine Pape, Silje L. Kaspersen, Lena Janita Skarshaug, Ellen Rabben Svedahl og Marlen Toch-Marquardt.

Hva skjer med informasjonen om deg?
Intervjuene vil tas opp på lydbånd og transkriberes. Du trenger ikke oppgi navn, hvilken kommune du bor i eller fastlegekontor i intervjuet. Datamaterialet lagres på sikre servere der kun prosjektdeltakerne har tilgang. Vi vil også registrere noen personopplysninger slik som kjønn, alder (ti-års kategorier), antall år som fastlege/helsesekretær (i kategorier) og spesialisering (for legene). I tillegg vil vi innhente noen opplysninger om legekontoret du jobber på (antall ansatte i ulike stillinger, oppgaver delegert til personell, rutiner for timebestilling, pasientsammensetning og håndtering av elektronisk kommunikasjon). Alle personopplysninger vil bli behandlet konfidensielt og kun prosjektgruppen vil ha tilgang til disse. Ditt navn, epostadresse eller telefonnummer vil lagres atskilt fra lydopptakene, og slettes ved prosjektslutt. Datainnsamlingen gjennomføres i 2017.

Studiens resultater vil bli publisert i flere forskningsrapporter. Rapportene vil ikke inneholde opplysninger som kan føres tilbake til deg eller din praksis. Prosjektet skal etter planen avsluttes **31.12.20**. Alle opptak og eventuelle personopplysninger vil da bli slettet.

Frivillig deltakelse
Det er frivillig å delta i studien, og du kan når som helst trekke ditt samtykke uten å oppgi noen grunn. Dersom du trekker deg, vil alle opplysninger om deg bli anonymisert, på lik linje med om du deltar.

Dersom du ønsker å delta eller har spørsmål til studien, ta kontakt med Silje L. Kaspersen – silje.l.kaspersen@sintef.no eller 95088303.

Forskningsprosjektet er tilrådd av Personvernombudet for forskning, referansenummer 54945

Samtykke til deltakelse i studien

På intervjutidspunktet vil vi samle inn skriftlig samtykke til deltakelse i forskningsprosjektet.

Jeg har mottatt informasjon om studien, og er villig til å delta

---
(Signert av prosjektdeltaker, dato)

Supplerende opplysninger om deltakere:
Følgende papirskjema fylles ut (utkrysning, noe tekst) av deltakerne selv i forbindelse med gruppeintervjuet. Opplysningene fra papirskjemaet vil sammenholdes og grovkategoriseres før de lagres elektronisk under et deltakernummer. Papirskjemaene vi deretter makuleres. Opplysningene vil via deltakernummeret kunne knyttes til stemmene på lydbånd, men kategoriseres slik at muligheten for å identifisere enkeltpersoner vil være liten. Dette deltakernummeret vil ikke kunne knyttes til kontaktinformasjon for deltakerne. Det vil ikke bli mulig å identifisere den enkelte lege når data blir publisert.

Legekontor		Id (vi fyller ut)

Kjønn
- ☐ Kvinne
- ☐ Mann

Alder
- ☐ 20-29 år
- ☐ 30-39 år
- ☐ 40-49 år
- ☐ 50-59 år
- ☐ 60-69 år (og eldre)

Antall år som helsesekretær/sykepleier
- ☐ Under 2 år
- ☐ 2-4 år
- ☐ 5-9 år
- ☐ 10-19 år
- ☐ 20 år eller mer

Utdanning
- ☐ Helsesekretær
- ☐ Sykepleier
- ☐ Annet — Hva:

Antall ansatte på legekontoret (totalt) - hvorav fastlegehjemler - hvorav helsesekretærer/ sykepleiere	
Gjennomsnittlig antall arbeidstimer per uke	
Antall pasienthenvendelser pr. dag (ca.)	
Spesielle karakteristika ved fastlegekontoret – sykehjemsleger, innvandrere/asylsøkere el.l.	

Opplysninger om legekontoret

Legekontor	
Modell for drift	
Antall ansatte	
Antall leger	
Antall støttepersonell (utdanning, Stillingsandel)	
Kommune	☐ Stor bykommune ☐ Mindre bykommune ☐ Kommune >5000 innbyggere ☐ Kommune <5000 innbyggere
Rutiner for timebestilling:	☐ Telefon ☐ SMS ☐ Internett ☐ «Time samme dag»
Ventetid	

ISBN 978-82-326-5430-7 (printed ver.)
ISBN 978-82-326-5841-1 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU
Norwegian University of
Science and Technology