Hazardous Drinking and Life Satisfaction in Norwegian Medical Doctors: Individual and Work-Related Predictors

A 15-Year Longitudinal and Nationwide Study

Javed Iqbal Mahmood, MD

Department of Behavioural Sciences in Medicine
Institute of Basic Medical Sciences
Faculty of Medicine
University of Oslo
2019
ACKNOWLEDGEMENTS

I am sincerely grateful to my supervisor Professor Reidar Tyssen for his guidance and research expertise, and for allowing me to conduct research in the University of Oslo’s Department of Behavioural Sciences in Medicine, Institute of Basic Medical Sciences, Faculty of Medicine. As a specialist in psychiatry and addiction medicine, beginning a research programme was challenging for me. However, Professor Tyssen provided a positive attitude and excellent supervision throughout this unforgettable research experience – from data analysis, to drafting my manuscript and thesis, to preparing presentations for international conferences. Without Professor Tyssen’s supervision and assistance, I would not have been able to complete my PhD. Professor Tyssen, thank you very much.

I also thank my co-supervisors Kjersti Støen Grotmol and Martin Tesli, for their time spent on our discussions and for their critical comments on my manuscript and thesis. Likewise, I am very grateful to the members of the Health Professional Research Group for their inspiring meetings. Gratitude also to Professors Emeritus Per Vaglum, Tore Gude and Arnstein Finset for their time spent on our discussions and for their valuable comments along the way. Special thanks to Professor Emeritus Torbjørn Moum, who assisted me with all of the statistical analyses in my papers; his knowledge of statistics and research methods made a great impression on me.

Thanks to all the other senior and fellow researchers in the Department of Behavioural Sciences in Medicine. Specific thanks to Professors Toril Dammen, Jarle Breivik and Veslemøy Wiese Ragna for their interesting discussions on different research topics during our lunch hours. I am particularly grateful to researchers Anneli Mellblom and Hanne Cathrine Lie for their guidance and support during my PhD training.

I thank my family. My sons Yousef and Ishaq and my wife Memoona have been understanding of my work and have never complained about my duty hours. I thank all of my brothers and sisters for their love and support. I am grateful for the love and guidance of my late parents, and for having taught me the value and importance of knowledge and education.

Finally, I thank God for all of the blessings he has bestowed upon my family and me.
ABSTRACT (SUMMARY)

Although previous studies have addressed Norwegian medical students’ and young doctors’ hazardous drinking, and possible mediators of this behaviour, we still lack representative and longitudinal studies on hazardous drinking and life satisfaction among doctors whose careers are more well-established. Studies have also shown that both individual- and work-based factors are associated with drinking and life satisfaction among doctors, yet prospective longitudinal studies are needed to identify possible risk factors and establish causality. Such information is important both to doctors personally and to the medical associations and organizational authorities that need to develop interventions to prevent doctors from developing alcohol problems, and to counteract doctors’ potential for life dissatisfaction. In addition, long-term longitudinal data from doctors may, to some extent, be generalized to other populations, especially with respect to individual predictors.

This thesis used data from the Longitudinal Study of Norwegian Medical Students and Doctors (NORDOC), in which surveys were administered to two nationwide cohorts of doctors: the Medical Student Cohort and the Young Doctor Cohort. The Medical Student Cohort (N = 421) commenced in 1993 with new medical students; the Young Doctor Cohort (N = 631) assessed students graduating from medical school in 1993 and 1994. These cohorts were spaced six years apart at baseline and both were subsequently surveyed five times, with the last survey administered in 2014 (NORDOC 6). Most previous NORDOC project papers have used longitudinal data from the Young Doctor Cohort only, whereas there are fewer publications from the Medical Student Cohort. The current study is the first to use NORDOC data from both cohorts, merged at each of four consistent time points in participants’ careers and analysed longitudinally.

This thesis consists of three research papers. Paper I examined possible risk factors present during medical school for later hazardous drinking. The prevalence of hazardous drinking declined from 14% at the end of medical school to 8% by year 10 post-graduation. About 23% of those who drank hazardlessly in medical school continued this behaviour during years 4 and 10 post-graduation. Drinking to cope with tension during medical school was found to be an important risk factor for hazardous drinking post-graduation, and males had a higher likelihood of hazardous drinking. Adjusted medical school predictors of hazardous drinking during post-graduation year 4 were male gender, use of alcohol to cope with tension, and hazardous drinking; older age, male gender and previous hazardous drinking were the adjusted predictors of hazardous drinking at year 10 post-graduation.
Paper II identified co-occurring work-related risk factors for hazardous drinking, while controlling for individual factors and mental distress (i.e., anxiety and depressive symptoms). Somewhat unexpectedly, there was no independent relationship between work stress, or other work-related factors, and hazardous drinking. Mental distress and life events were independently related to hazardous drinking in a repeated-measures model, whereas drinking to cope (i.e., during medical school) was an independent predictor of hazardous drinking. Men were at higher risk of hazardous drinking.

Paper III examined trends in life satisfaction over the 15 years following graduation, and whether work stress had an independent impact on overall life satisfaction among doctors during this period (after controlling for individual and lifestyle confounders, such as drinking behaviours). Doctors’ mean life satisfaction ratings did not change across the three observational follow-up periods. There were no significant gender differences in level of life satisfaction. The work-related factors that were independently associated with higher life satisfaction were low work-home stress, low perceived job demands and high colleague support. Both hazardous drinking and drinking to cope were independently related to life satisfaction. In addition, those who experienced a decrease in life satisfaction showed higher levels of neuroticism personality trait and lower colleague support.

Consistent with studies of other populations, we found that individual factors, mental distress and life stress were the most important risk factors for hazardous drinking among doctors. With respect to their work-related factors, we found no link with hazardous drinking, but found an association between these factors and doctors’ life satisfaction. This was true for work–home stress, perceived job demands and colleague support. Low colleague support was another important predictor of doctors’ reduced life satisfaction during the follow-up years. These findings support the notion that doctors’ work-related factors might be improved through organizational-level interventions, including a more supportive work environment, towards increasing their life satisfaction. In addition, preventive measures should be taken to reduce medical students’ and doctors’ drinking to cope with tension and mental distress, particularly for men who are at higher risk of hazardous drinking.
LIST OF PAPERS


ABBREVIATIONS

NORDOC = The Longitudinal Study of Norwegian Medical Students and Doctors
BCI = Basic Character Inventory
NMA = Norwegian Medical Association
SOP = Nursing and Pension Scheme for Doctors
QPS Nordic = General Nordic Questionnaire for Psychological and Social Factors at Work
LG = Logistic regression
GEE = Generalized estimating equation
LMM = Linear mixed model
AUDIT = Alcohol Use Disorder Identification Test
CAGE = Cut down drinking, Annoyed by criticism, Guilty feelings, and Eye opener
SPSS = Statistical Package for the Social Sciences
OR = Odds ratio
CI = Confidence interval
P = P-value
β = Unstandardized regression coefficients
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ............................................................................................................. 2
LIST OF PAPERS .......................................................................................................................... 5
ABBREVIATIONS .......................................................................................................................... 6

1. INTRODUCTION ....................................................................................................................... 9
   1.1 What is this thesis about? ........................................................................................................ 9
   1.2 How do we define hazardous drinking? ............................................................................... 9
   1.3 How do we define life satisfaction? .................................................................................... 11
   1.4 Why study doctors' hazardous drinking? ........................................................................... 12
   1.5 Previous longitudinal studies ............................................................................................ 13
   1.6 Why study doctors' life satisfaction? .................................................................................. 14
   1.7 Prevalence of alcohol abuse among doctors ...................................................................... 16
   1.8 Predictors of hazardous drinking in longitudinal studies .................................................... 18
   1.9 Factors associated with hazardous drinking in cross-sectional studies ............................... 18
   1.10 Life satisfaction ............................................................................................................... 23
   1.11 Factors associated with life satisfaction ........................................................................... 25

2. AIM OF THE THESIS ............................................................................................................... 29
   2.1 Research questions ............................................................................................................. 29

3. MATERIALS AND METHODS ................................................................................................. 30
   3.1 Historical background of the NORDOC project: The Longitudinal Study of Norwegian Medical Students and Doctors ................................................................................. 30
   3.2 Study design ....................................................................................................................... 31
   3.3 Sample ................................................................................................................................ 31
   3.4 How representative of Norwegian medical doctors is our sample? .................................... 33
   3.5 Dependent variables .......................................................................................................... 35
      3.5.1 Hazardous drinking (single-item measurement) .......................................................... 35
      3.5.2 Hazardous drinking (nine-item version) .................................................................... 37
      3.5.3 Life satisfaction ........................................................................................................... 38
   3.6 Independent variables ........................................................................................................ 38
      3.6.1 Socio-demographic variables (age, gender, married/cohabitating) ......................... 38
   3.7 Other predictor variables .................................................................................................. 39
      3.7.1 Having children or not ............................................................................................... 39
      3.7.2 Religious activity ......................................................................................................... 39
      3.7.3 Perceived social support ............................................................................................ 39
      3.7.4 Physical activity ......................................................................................................... 40
      3.7.5 History of parental alcohol problems ....................................................................... 40
      3.7.6 Cohort ......................................................................................................................... 40
      3.7.7 Personality traits ....................................................................................................... 40
      3.7.8 Use of alcohol to cope with tension ......................................................................... 42
      3.7.9 Life events ................................................................................................................ 42
      3.7.10 Mental distress ......................................................................................................... 43
      3.7.11 Work-related variables ............................................................................................. 43
   3.8 Statistical analyses .............................................................................................................. 48
   3.9 Discussion of methodological issues .................................................................................... 50
      3.9.1 Study design .............................................................................................................. 50
      3.9.2 Selection bias ............................................................................................................ 51
3.9.3 Reporting bias ................................................................................................. 52
3.9.4 Missing data .................................................................................................. 52
3.9.5 Psychometric considerations ......................................................................... 53
3.9.6 External validity ............................................................................................. 55
3.9.7 Ethics ............................................................................................................. 56

4. SUMMARY OF PAPERS .................................................................................. 57

4.1 Paper I .............................................................................................................. 57
4.2 Paper II ............................................................................................................. 58
4.3 Paper III ........................................................................................................... 59
4.4 Additional findings (not previously published) .............................................. 60

5. GENERAL DISCUSSION OF THE RESULTS ........................................... 62

5.1 Prevalence of hazardous drinking among doctors as a group .................... 62
5.2 Stability of hazardous drinking among doctors as a group ............................. 63
5.3 Individual predictors of hazardous drinking .................................................. 64
5.4 Additional findings as predictors of enduring hazardous drinking ............... 67
5.5 Independent factors related to hazardous drinking ......................................... 67
5.6 Level of life satisfaction among Norwegian doctors (longitudinal sample) .... 70
5.7 Independent work-related factors associated with life satisfaction ............... 70
5.8 Other individual predictors or factors linked to life satisfaction ..................... 73
5.9 Predictors of a decline in life satisfaction ....................................................... 75
5.10 Drinking behaviours independently linked to life satisfaction ..................... 76
5.11 Gender differences in predictors of hazardous drinking ............................... 76
5.12 Gender differences in predictors of life satisfaction ....................................... 77
5.13 Cohort effects .................................................................................................. 78
5.14 Implications for medical school and clinical practice .................................... 78
5.15 Implications for future research ................................................................. 80
5.16 Conclusions ................................................................................................... 81

REFERENCES ..................................................................................................... 82
1. **INTRODUCTION**

1.1 **What is this thesis about?**

This thesis addresses drinking patterns and life satisfaction among Norwegian doctors, who were followed from their final year of medical school until 15 years post-graduation. During this period, from their early training to their having become experienced doctors, we prospectively studied their hazardous drinking and life satisfaction, to identify possible predictors of, and coexisting associations with, these dependent variables. We particularly focused on individual factors, such as vulnerable personality traits, drinking to cope with tension and mental distress (i.e., anxiety and depressive symptoms). In addition, we studied co-occurring work-related factors that may be associated with both hazardous drinking and life satisfaction during this 15-year span. Furthermore, drinking behaviours (both hazardous drinking and drinking to cope with tension) were among the many independent variables included in repeated-measures models of life satisfaction. This thesis was designed to investigate these issues, including the questions: ‘Is it possible to identify risk factors present during medical school for hazardous drinking and life satisfaction 15 years post-graduation?’ and ‘Is it possible to identify work-related factors related to hazardous drinking and life satisfaction during these 15 years?’ The answers to these questions may facilitate the development of preventive measures during medical school, and throughout doctors’ early-to-mid-career period, to reduce their risks of developing serious alcohol problems and poor life satisfaction.

1.2 **How do we define hazardous drinking?**

At present, there is no consensus on the definition of hazardous drinking (i.e., binge drinking, problematic drinking). The National Institute on Alcohol Abuse and Alcoholism has defined hazardous drinking as ‘a pattern of drinking that brings blood alcohol concentration levels to 0.08 g/dl. This typically occurs after 5 drinks for men and 4 drinks for women—in about 2 hours’ (1). The US Substance Abuse and Mental Health Services Administration defines hazardous drinking as ‘drinking ≥ 5 alcoholic drinks for males or ≥ 4 alcoholic drinks for females on the same occasion, that is, at the same time or within a couple of hours of each other on at least one day in the past month’ (2). The Norwegian Directorate of Health recommends not drinking more than 10 g of pure alcohol/day for women and 20 g/day for men (3), which is consistent with the US Department of Health and Human Services and the
US Department of Agriculture 2015–2020 Dietary Guidelines for Americans (4). Figure 1 shows various patterns of alcohol use, from abstinence to low-risk use, risky use, problem drinking, harmful use (i.e., alcohol abuse), and alcoholism, which is the most severe form of alcohol dependence (5). The green apex of this pyramid represents diagnosable alcohol use disorders with definite pathology. In this thesis, hazardous drinking corresponds to problem drinking (i.e., unhealthy, albeit non-disordered, drinking).

Low-risk use is the most common alcohol use pattern among the general population. This is also known as social drinking, in which alcohol is consumed in very small, safe amounts. In contrast, risky use for men aged < 65 years is > 4 drinks/occasion, or > 14 drinks/week for men; for women these values are > 3 drinks/occasion or > 7 drinks/week. For men and women aged ≥ 65 years this becomes one drink/day (6). A lengthy review recently published in the Lancet recommends lowering these thresholds based on evidence of risk for heart disease and other disorders (7). The review’s primary finding was a positive, curvilinear association between all-cause mortality and level of alcohol consumption; those consuming around or under 100 g/week (about 5–6 UK standard glasses of wine or pints of beer/week) were at lowest risk and binge drinkers were at highest risk.

Importantly, although hazardous drinking is not yet considered a disorder (i.e., does not constitute alcohol abuse or dependency), hazardous drinking over time increases the risk of alcohol dependency (8). It is therefore important to identify hazardous drinking early, to prevent individuals from developing alcohol-related disorders. The research group conducting the Longitudinal Study of Norwegian Medical Students and Doctors (NORDOC) has in previous papers used a single-item measure of hazardous drinking, defined as ‘drinking a total of 60 g of ethanol (approximately five alcoholic units) or more in one session at least 2–3 times per month during the last year’ (9-11). This single item measuring hazardous drinking has not been validated except in studies of US college students, which linked this drinking level to harmful psychosocial consequences (12). Therefore, in the present study, we also validated this item for use with practising doctors in their late 30s and early 40s, with respect to somatic health problems, which are related to increased alcohol use (13-15). We also used a longer (nine-item) version of the Alcohol Use Disorder Identification Test (AUDIT), which has likewise not been well validated to assess hazardous drinking. Thus, we validated both our single-item measure and the nine-item version of the AUDIT with respect to somatic health complaints, which are known to be associated with drinking.
Fig. 1. Spectrum of alcohol use from abstinence (at the base) to alcohol dependence (at the apex).

1.3 How do we define life satisfaction?

With respect to the concept of wellness, different terminologies have been used to define physician wellness (e.g., ‘life satisfaction’, ‘well-being’, ‘happiness’, ‘quality of life’, ‘health’) (16), all of which are interrelated, although attempts have been made to define some of these from the perspective of positive psychology (17). Positive psychology, a branch of psychology that has developed over the past 20 years, asserts that wellness is about resilience, strengths and the factors that make people thrive (18). Positive psychology focuses on prevention and how human strengths such as hope, honesty, courage, wisdom, optimism and insight may buffer against stress and mental disorders (18). A recent systematic review of physician wellness showed that the majority (86%) of published papers have assessed physician wellness without providing a clear definition of the construct (16). These authors proposed the following definition: ‘Physician wellness (well-being) is defined by quality of life, which includes the absence of ill-being and the presence of positive physical, mental,
social, and integrated well-being experienced in connection with activities and environments that allow physicians to develop their full potentials across personal and work-life domains’. Their definition accounts for the domains of: 1) negative (ill-being) and positive (well-being); 2) physical, mental, social and integrated well-being; and 3) work life and personal (non-work, general) life. The NORDOC study group has previously published studies on life satisfaction among medical students (19) and doctors (20) using a single-item measure of this construct. In this study, we used a more reliable three-item measure and applied repeated-measures analyses of the three follow-up periods to increase the reliability of our findings.

1.4 Why study doctors’ hazardous drinking?
Alcohol is the most commonly abused substance among doctors who seek addiction treatment (21-23). Hazardous drinking is associated with increased risk to self and other physical and psychological harms (24) and rates of self-reported medical error are higher among doctors who have alcohol abuse or dependency problems (25, 26). Alcohol and drug abuse are among the leading causes of disciplinary proceedings against doctors by state licensing authorities (27). Doctors are educated about the negative health-related consequences of alcohol and might therefore be presumed to drink less compared with the general population; indeed, doctors smoke less compared with others (28). Doctors may also serve as role models in terms of their alcohol consumption behaviours, and their own drinking patterns may influence the advice they give their patients regarding alcohol use (29). Studies have shown that doctors as a group seem to drink at about the same level as comparable groups in the general population (30, 31), although one US study reported that doctors drink more compared with the general population (25). Unfortunately, there is limited updated information about doctors’ drinking patterns in Norway, a temperance country with strict regulation on the sale of alcoholic beverages. We lack studies on doctors’ drinking in such cultural context. This work may also indicate whether the Norwegian context influences alcohol misuse in unique ways compared with other countries. This prospective assessment was designed to clarify the importance of work, individual and other factors in accounting for doctors’ drinking patterns. We were also interested in determining whether doctors’ drinking behaviours are associated with their well-being over a relatively long-term follow-up period. The study may provide insight into how to improve work-related factors, and possible interventions at the work places in the future.
1.5 Previous longitudinal studies

The majority of studies on doctors’ drinking patterns have been cross-sectional (32-37); to our knowledge, only four longitudinal studies have addressed this topic (30, 38-40). These four studies are described below.

The first longitudinal study was published in 1990 by Moore et al., who analysed data from an ongoing, prospective study (the Johns Hopkins Precursors Study) of 1,014 male medical students who graduated during 1948–1964 from the Johns Hopkins School of Medicine (38). The study participants underwent extensive assessments to determine young doctors’ characteristics that may affect their later health. Data collected during medical school were examined as possible predictors (or precursors) of alcohol abuse during subsequent mid-life (ages 52–68 years). Mid-life alcohol abuse was reported by 12.9% of this sample and independent predictors were non-Jewish ancestry, lack of religious affiliation, smoking, regular use of alcohol, anxiety or anger as a reaction to stress, non-social use of alcohol and a history of maternal mental illness or alcoholism. However, sleeping habits, substance abuse (sedatives or amphetamines), interest in athletics/hobbies and parental relationship were not significantly associated with mid-life alcohol abuse.

The second longitudinal study, by Newbury-Birch et al. in 2001, assessed drinking and illicit drug use in a cohort of medical students surveyed during years 2 and 5 of medical school, and then one year later when they were pre-registration house officers (39). Mean units/week of alcohol consumed increased significantly over the three observation periods, from 15.2 (year 2) to 16 (year 5) to 18.8 (one year after working as pre-registration house officers). However, these data were analysed as three independent samples, not longitudinally or as repeated measures.

The third longitudinal study was a 23-year prospective study, including three follow-up time points, of 12,000 male British doctors aged 48–78 years in 1978. In 2005, Doll and colleagues examined mortality in relation to alcohol consumption in this older population (40). Mean alcohol consumption in this sample was 2–3 units/day and causes of death known to be increased by alcohol use accounted for only 5% of their deaths (1% liver disease, 2% cancer of the mouth, pharynx, larynx, or oesophagus, and 2% external causes of death); however, the rates of these deaths were significantly higher among those consuming > 2 units/day. For other causes of death, and for the all-cause aggregate, a U-shaped dose–response relationship showed higher risks for non-drinkers and those reporting an average of > 4 units a day (28/week), compared with those reporting intermediate consumption levels.
This dose–response relationship was not significant for cancer, but was highly significant for vascular disease, respiratory disease and all-cause mortality.

The fourth longitudinal study was by Rosta and Aasland in 2013, who used data from two waves (2000 and 2010) of a representative cohort study of 682 Norwegian doctors (30). These authors investigated changes in patterns and consequences of alcohol use among Norwegian doctors, and showed that drinking patterns had changed between 2000 and 2010 towards more moderate alcohol consumption and fewer negative alcohol-related consequences. Doctors drank more, but with fewer episodes of heavy drinking and fewer alcohol-related problems. The AUDIT score for dependency symptoms was unchanged across the decade.

The second and fourth studies described above were relatively short and included fewer follow-up periods, addressing doctors’ drinking patterns (or alcohol consumption) but not assessing predictors of hazardous drinking (30, 39). The third study examined mortality in relation to alcohol consumption among older male British doctors (40). The first study by Moore et al. (38) was the only one that examined predictors of later alcohol abuse; although only male participants were included, several independent predictors of mid-life alcohol abuse were identified.

In a six-year prospective study, our research group discovered that a low level of control trait (or conscientiousness) and drinking to cope were predictors of hazardous drinking among medical students (10). In a separate six-year longitudinal NORDOC study, Grotmol et al. (11) found that alcohol expectancy was a predictor of later hazardous drinking among doctors, an effect not mediated by drinking to cope with tension. No previous NORDOC study has examined long-term predictors of hazardous drinking among doctors after they leave medical school; nor has any previous study validated our single-item hazardous drinking measure with respect to alcohol-related physical health problems (13-15). Furthermore, NORDOC researchers have previously used longitudinal data from each of the cohorts separately. In the present study, we used longitudinal data from both cohorts, which we merged at each of the four comparable career time points. This approach provided a larger, more robust sample, which increased our statistical power.

1.6 Why study doctors’ life satisfaction?

Life satisfaction may be important to doctors’ mental health and general well-being. Doctors are at increased risk for suicide, burn-out and depressive symptoms. The first postgraduate
years are considered particularly stressful (41-44) and there is some evidence of a co-occurrence of depression, alcohol/drug abuse and suicide among both doctors (45, 46) and within the general population (47-49). These comorbidities were the subject of a thesis on mental health problems in Norwegian medical students and doctors (50). A meta-analysis of 25 studies (mainly from northern Europe and North America) on doctor suicide reported aggregate suicide rate ratios for male and female doctors to be 1.41 and 2.27, respectively, compared with the general population (51). This means that the rate of death by suicide among male doctors is about 40% higher than among other men; for female doctors, the suicide rate is about 130% higher than among other women. In Norway, doctors show higher suicide rates compared with both other university graduates and the general population; this is true for both male (43/100,000 person-years) and female (26.1/100,000 person-years) doctors (52). In addition, suicide rates have increased from 1960 to 2000, after adjusting for age, among both physicians and other university graduates. Increased suicide prevalence is the most consistent mental health-related risk among doctors worldwide. Suicide rates among the general Norwegian population have been quite stable since 2000 (53), although it is unknown whether this also applies to Norwegian doctors. Therefore, it is important to study factors affecting doctors’ well-being and life satisfaction. One cross-sectional study in the USA reported that alcohol abuse or dependence is associated with depression, suicidal ideation, burn-out, lower career satisfaction and lower quality of life among doctors (25).

Based on these cumulative data, it is unknown whether hazardous drinking, which has not yet been defined as alcohol use disorder, may also be linked to lower life satisfaction among doctors. Furthermore, drinking to cope with tension may be another indicator of hazardous drinking. Therefore, drinking behaviours such as hazardous drinking and drinking to cope with tension should be included in longitudinal studies of doctors’ life satisfaction. A major aim of this study was to identify work-related factors that are independently linked to doctors’ life satisfaction. Combining the two NORDOC cohorts allowed us the unique opportunity to use large predictor models of both individual- and work-related factors that may impact life satisfaction over a relatively long period in doctors’ careers. These are also potentially modifiable factors that could be targeted through workplace interventions to counteract negative mental health consequences such as burn-out and depression.
1.7 Prevalence of alcohol abuse among doctors

The fact that doctors’ alcohol use is similar to that among the general population is puzzling, since drinking problems are fairly widespread and doctors tend to adopt healthier lifestyle habits with respect to smoking (23, 54) and physical activity (55-57). However, as most studies of doctors’ alcohol use have been from North America, less is known about the situation in European countries. Since Norway and Sweden are relatively restricted with respect to the sale of alcoholic beverages, and have a more conservative drinking culture compared with other European and North American countries, alcohol use data from other countries may not generalise to Norway and Sweden.

One cross-sectional study of US surgeons showed that 15.4% (14% for males, 26% for females) had a score on AUDIT-C consistent with alcohol abuse or dependency (26), while another cross-sectional study in the USA reported that 13% of male and 21% of female physicians met diagnostic criteria for alcohol abuse or dependency during the past 12 months (25). However, because the response rates for these two studies were 28.7% and 26.7%, respectively, their samples may have been vulnerable to selection bias and may not be representative of US physicians. Furthermore, there are no recent data among US physicians with respect to hazardous drinking or alcohol use disorders.

Contrary to these US findings, a 2008 cross-sectional study by Rosta showed lower rates of binge drinking (at least monthly) among hospital doctors in Germany compared with the general population (13.5% vs. 38%, respectively), although over 20% of doctors in surgical specialties reported hazardous drinking (58).

A high-quality review of representative samples of doctors by Rosta in 2005 reported overall problem-related drinking, from heavy/hazardous drinking (12–16%) to misuse and dependence (6–8%), among several population based samples of doctors (59). Rosta concluded that representative samples are also needed to identify risk factors related to problem drinking among doctors, and that drinking problems should be studied in the context of health, lifestyle and work-related factors to develop preventive measures for doctors.

In 2005, Rosta and Aasland reported that female Norwegian surgeons had a higher rate of hazardous drinking compared with other female physicians (18% vs. 7.6%, respectively) and that male surgeons engaged in more hazardous drinking than other male physicians (60). The 86% (1385/1616) response rate in this study was high, although among the 226 surgeons studied, only 56 were female, reflecting a possible selection bias. The proportion of Norwegian doctors who engage in hazardous drinking increased from 12% in
1993 to 15% in 2000, using an AUDIT cut-off score \( \geq 6 \) (61), but there was a decline from 10.7% in 2000 to 8.2% in 2010 using an AUDIT cut-off score \( > 8 \) (30). In a study from the 1980s, female doctors in Norway were reported to drink more often than other academics (32). In one Finnish study, about 16% of doctors reported high alcohol consumption \((> 200 g/week)\) (33), although to our knowledge no recent report is available for Finnish doctors. A recent nationwide, cross-sectional Danish study reported that 18.3% of doctors drink hazardousily (AUDIT cut-off score \( > 8 \)) (62); the same group also reported that the highest proportion of risky alcohol use (24%) was among internal medicine and emergency medicine doctors, while the lowest proportion (16%) was among general practitioners (36). Based on these data, the level of hazardous drinking (using AUDIT cut-off \( > 8 \)) among Danish doctors is much higher than the values reported above for Norwegian doctors: 18.3% vs. 8.2%, respectively. The Danish study also had a relatively low response rate (49%), which may have caused some selection bias. Nevertheless, the same proportion (18%) of hazardous drinking was found among medical specialists in a Belgian study (20.7% among men and 15% among women) (35). This study also revealed that medical specialists engaged in more hazardous drinking compared with the general population (18% vs. 10%, respectively), although the specialists’ rate was consistent with that among female Norwegian surgeons described above.

In 2013, Wurst et al. (63) reported survey data from Salzburg, Austria, for which the study aim was to identify precise prevalence rates of alcohol problems among doctors. This was accomplished by correcting for false positives and negatives using a published correction formula based on general population data and instrument sensitivity and specificity values. Using the full AUDIT, 19.6% of female doctors and 48% of male doctors screened positive using a cut-off score of \( \geq 5 \), while 9.3% of female doctors and 16.8% of male doctors screened positive using a cut-off score of \( \geq 8 \). After correction based on the AUDIT sensitivity and specificity (cut-off score \( > 5 \)), the estimated prevalence rates for female doctors declined from 19.6% to 6.3% and those for male doctors from 48% to 15.5%. However, this study had a response rate of only 18.4%, which is quite low.

In summary, these cumulative findings show that doctors engage in hazardous drinking to the same extent as the general population, although direct comparisons are difficult because different instruments and/or cut-off values have been used for defining hazardous drinking. Recent studies from the USA have been limited by low response rates and under-reporting is always a problem in studies of drinking behaviours. Furthermore, the
US-based studies focused on later-stage problem drinking, such as alcohol abuse or dependency (i.e., alcohol use disorders), whereas the European studies focused on early-stage problem drinking (i.e., harmful or hazardous drinking). Interestingly, the level of hazardous drinking among Danish and Belgian doctors appears to be at least double that among Norwegian doctors, which likely reflects the different drinking cultures in these countries. Compared with the large number of studies addressing doctors’ stress and burn-out, there are quite few that analyse their drinking behaviours. Thus, this was the focus of this thesis.

1.8 Predictors of hazardous drinking in longitudinal studies

To our knowledge, only four prospective longitudinal studies have assessed hazardous drinking among doctors, the details of which were described in section 1.5 (30, 38-40). The only prospective study to examine possible predictors of doctors’ subsequent alcohol abuse during mid-life was by Moore et al. (38). They identified several variables related to hazardous drinking: non-Jewish ancestry, a history of problems caused by drinking, smoking, a history of maternal mental illness or alcoholism, non-social use of alcohol and anxiety as a reaction to stressful situations. These variables explained 23% of the variance between those who subsequently abuse alcohol in mid-life, and those who do not.

1.9 Factors associated with hazardous drinking in cross-sectional studies

A large number of cross-sectional studies of doctors have identified several statistical predictors (or concurrent factors) associated with alcohol consumption. Because such predictors of hazardous drinking have been barely studied longitudinally, in representative samples, we aimed to identify factors that prospectively predicted hazardous drinking among doctors post-graduation. Another major study aim was to identify independent work-related factors associated with hazardous drinking. Before discussing these factors, we will describe individual factors to be controlled for in our prediction models.

Age. Several cross-sectional studies of doctors have shown healthier drinking habits among younger doctors compared with their older counterparts (34, 35, 58). Literature reviews have shown that the highest prevalence of problematic drinking is among young adults and college students (64, 65). However, one cross-sectional study of the general Finnish population showed hazardous drinking to be more prevalent among those aged 40–49 years (66).

An increased awareness of, and emphasis on, the dangers or consequences of alcohol-related harm may contribute to lower alcohol consumption among younger doctors. Another
reason for this may be having a family and taking on the responsibilities of parenthood. Our 
research group has shown that having children is associated with reduced hazardous drinking 
among medical students (9), although updated studies of hazardous drinking in representative 
samples of doctors at the mid-career stage are needed.

**Gender.** Several cross-sectional studies have consistently shown that male doctors engage in 
more hazardous drinking compared with female doctors (36, 37, 58, 63), which is a gender 
distribution reflected in the general population (67-70). This trend was also shown in a cross-
sectional study of veterinary surgeons in the UK (71).

With respect to hazardous drinking among medical students, several European cross-
sectional (72-74) and prospective longitudinal (11, 39) studies of medical students have 
shown that male students drink more hazardingously compared with female medical students. 
Similar results have been reported in US cross-sectional (75) and prospective cohort (76) 
studies of medical students, other college students (77, 78) and the general population (79).

**Cohort.** Since the two cohorts in the present study were initiated six years apart, combining 
them made it possible to evaluate changes in drinking trends over recent years among 
younger participants. For instance, we know that the annual consumption of pure ethanol per 
inhabitant in Norway increased by about 16% (about one litre) from 2000 to 2010. During the 
same period, hospital admissions for alcohol-related disorders increased by 48% (80). Thus, 
it is important that we control for a cohort effect in our statistical model.

**Not having children.** Our research group previously identified not having children as an 
independent predictor of hazardous drinking among medical students, which may be 
explained by the lack of family responsibilities (9). This is consistent with two cross-sectional 
 studies from the USA showing that having children is associated with a lower likelihood of 
 alcohol abuse or dependence (25, 26).

**No religious activity.** One longitudinal study showed that a lack of religious activity is 
 linked to increased drinking among doctors (38), a finding similar to those from cross-
sectional studies of both medical students (9, 72) and the general population (81-84). The role 
of negative attitudes towards drinking related to religious activity was therefore controlled for 
in our prediction model.
History of parental alcohol problems. A family history of alcoholism is a known risk factor for development of alcohol problems in both the general population (85, 86) and among doctors (31, 87). This has also been confirmed in a longitudinal study of doctors (38).

Personality traits. To our knowledge, few prospective studies have evaluated the relationships between personality and drinking among doctors. A prospective study by Richman et al. showed that the interaction between experiencing workplace abuse and personality vulnerability (narcissism) predicted doctors’ drinking problems (88). Brooke et al. surveyed 144 doctors who had received substance misuse treatment to find that the most frequent cause of developing a substance use disorder was personality difficulty (52.8%) (89).

A meta-analysis of eight adult cohort studies from the USA, the UK, Germany and Australia (N = 72,949) showed increased alcohol consumption to be more common among extraverts and those with low conscientiousness (i.e., not being organized/orderly/capable of fulfilling tasks) whereas those with high agreeableness (being generous, honest, sympathetic) and low openness (curious, engaging in fantasy, day dreaming) had an increased odds of reduced alcohol consumption and preferred abstinence (90). The NORDOC research group previously found that the personality trait of low conscientiousness, measured at the beginning of medical school, predicted hazardous drinking six years later (10). Such traits have never been assessed for their ability to predict drinking patterns in doctors after they leave medical school; thus, assessing the predictive validity of personality traits with respect to drinking during doctors’ careers was one goal of using the NORDOC cohorts for this thesis study.

Use of alcohol to cope with tension. Drinking to cope with tension has previously been validated as a predictor of hazardous drinking in medical students by the NORDOC group, including prospectively over six years (9, 10). A cross-sectional British study showed that about 26% of hospital consultants drink alcohol to cope with work stress, and also have an increased risk of psychiatric morbidity (91). Such drinking behaviour should be further studied among doctors. To that end, among the theories and models on the aetiology of alcoholism, we particularly focused on two: the tension reduction hypothesis (92) and the self-medication hypothesis (93, 94). Both models involve negative affect regulation.
The tension reduction hypothesis states that alcohol is believed to reduce tension and that individuals learn to drink alcohol because of its tension-reducing effects (i.e., reinforcement). This theory is related to expectancy theory in the sense that individuals drink because they believe alcohol will reduce their negative emotions (95). The NORDOC group previously found that expectancy, which relates to cognitive factors, is an independent risk factor for, and prospective predictor of, hazardous drinking (11). In the present study, we included use of alcohol to cope as a predictive risk factor for later hazardous drinking, with the hypothesis that drinking to cope with tension would be an independent predictor of hazardous drinking after leaving medical school. However, drinking to cope is also a drinking behaviour, whereas the expectancy theory refers to drinking beliefs (i.e., cognitions about alcohol).

The self-medication hypothesis was developed based on clinical work with patients who have co-occurring psychiatric illness and substance use disorder. This comorbidity is also a risk factor for relapse in clinical samples of doctors-as-patients (96). According to this hypothesis, substances are abused to relieve psychological pain/suffering, with a specific link between the patient’s substance preference and specific, intolerable symptoms that they want to reduce or alleviate. Successfully treating co-occurring psychiatric disorders, such as depression, may alleviate substance use disorders (97). Use of alcohol as a tension-relieving ‘drug’ is thus consistent with this hypothesis.

Life events (during the past year). The influence of negative life events on drinking among doctors has not been well studied. Within the general population, Veenstra et al. reviewed the literature on the relationship between life events and alcohol use. Four longitudinal studies have shown that life events related to a spouse, friends or relatives lead to an increase in alcohol use, whereas health-related life events and financial problems precede a decrease in alcohol use (98). A longitudinal birth cohort study in which participants were followed to age 30 years showed that those with the highest exposure to stressful life events had an odds of alcohol abuse/dependency 2.24 times higher than those at the lowest level of life event exposure (99).

Mental distress (past two weeks). Several studies have shown that doctors have relatively high levels of anxiety and depressive symptoms when they are in medical school, during their internship and during the first postgraduate years (44, 100). A cross-sectional study in
Canada reported that about 42% of medical students and residents were affected by anxiety and depressive symptoms (101). Research clearly shows that some doctors face mental health problems as they progress through their careers, which may increase their risk of developing a substance use disorder. To our knowledge, no representative studies of doctors have used diagnostic interviews for mental health disorders, thus it is unclear whether levels of these may be higher among medical students and doctors compared with other student and academic groups. A survey of US medical students, residents and fellows (i.e., house staff) found that ‘drinking too much’, as a measure of potentially risky drinking, was linked to several mental health problems, including severe depressive symptoms and impairment, past suicide attempts and current suicidal ideation, intense affective states and other substance use (102). Therefore, co-occurring mental distress and alcohol use may be important contributors to doctors’ development of alcohol problems, as is the case in the general population. Among the few studies using clinical samples of doctors, a Spanish study found that the most common dual diagnosis among doctors was comorbid alcohol use disorder and affective disorders (46).

**Work stress.** To our knowledge, there is little empirical evidence of a direct relationship between work/occupational stress and hazardous drinking or alcohol use among either doctors (103, 104) or the general population (based on a meta-analysis) (105-107). However, work stress may be linked to job dissatisfaction (108, 109), depressive symptoms (110) and burn-out (111, 112), which may lead to heavy drinking, as shown among Finnish doctors (33). However, this Finnish study was conducted 30 years ago. Recent cross-sectional studies of US physicians have shown that alcohol abuse or dependence is associated with burn-out, depressive symptoms, suicidal ideation, lower quality of life and lower career satisfaction (25, 26), although these studies cannot deduce direction or causation. Sleep deprivation and negative perception of work environment were associated with alcohol and/or drug abuse in a French national survey of doctors (113). Evidence from longitudinal studies on the link between work-related stress and drinking among doctors is lacking, particularly regarding the types of doctors’ work pressures that may be linked to their drinking (e.g., emotionally demanding patient work, time pressures, fear of mistakes and litigation, work–home stress) (114). Work-related factors may also impact male and female doctors differently, thus gender-based interaction effects should be controlled for in prediction models (115). A previous NORDOC study showed that among young doctors, work stress is related to mental
health problems during their early careers and that the most important among these stress factors are emotional pressure and patient demands (116). Nevertheless, other work pressures should also be included in prediction models, such as psychosocial work-related factors.

A vast amount of research has been conducted over the past three decades, including the development of various theoretical concepts/models of psychosocial work-related factors to explain the effects of work stress on health in the general population. In particular, Karasek and Theorell’s demand–control model (117) has received a great deal of attention. Thus, in this NORDOC project, we used perceived job demands and autonomy variables resembling Karasek’s demand–control variables in the General Nordic Questionnaire for Psychological and Social Factors at Work (QPS Nordic) (118).

The demand–control model identifies high work stress based on job tasks characterized by high quantitative demands in combination with low control. A low level of control (or decision latitude) manifests in two ways: 1) lack of decision authority over one’s own tasks, and 2) a low level of skill utilization. This model was later expanded to include a dimension of social support at work, which may act as a buffer on the relationship between high demand and low control (119). We therefore included colleague support as a potentially important measure among the work-related factors that may be linked to hazardous drinking. According to this hypothesis, we would expect perceived job demands and autonomy to impact doctors’ drinking behaviours.

1.10 Life satisfaction

In addition to hazardous drinking, life satisfaction among Norwegian doctors was the second primary outcome analysed in this thesis study. In contrast to studies of doctors’ stress and poor mental health, few studies have addressed the relationship with positive psychology and life satisfaction (120-122). As noted, life satisfaction may be important with respect to a doctor’s own mental health and, consequently, abuse of alcohol. Furthermore, it is unknown whether our measure of hazardous drinking is linked to lower life satisfaction among doctors. A growing body of literature addresses dissatisfaction among doctors (123-127). An online survey of 24,000 US doctors in 2012 showed that only 54% would choose medicine again as a career, a decline from 69% in 2011 (128). A national survey of general practitioners in England showed an increase from 14% (in 1998) to 22% (in 2001) on intention to leave medicine within the next five years (129). A previous NORDOC report showed lower life satisfaction among Norwegian doctors during their first postgraduate years compared with
that among the general population, and that work stress is an important factor contributing to lower life satisfaction during doctors’ early careers (20). However, we do not know exactly what type of work stress and/or work-related factors affect doctors’ well-being beyond their initial postgraduate years. To our knowledge, we lack representative and longitudinal prospective studies that identify the types of work stress and work-related factors that are associated with lower life satisfaction among doctors; understanding these is important for developing organizational interventions to increase overall well-being among doctors (130). Thus, such studies are highly warranted (131), including the present thesis.

Recently, there seems to have been a shift away from focusing exclusively on doctors’ negative mental health. Eckleberry-Hunt and colleagues argue: ‘We must move beyond the pathological focus upon physician burnout and begin a conversation about what makes a physician well’ (132). A 2015 article published in the Lancet identified three aspects of subjective well-being (133): 1) evaluative well-being (or life satisfaction), which considers how satisfied people are with their lives; 2) hedonic well-being (everyday feelings or moods such as happiness, sadness and anger); and 3) eudemonic well-being (judgement about the meaning and purpose of life). In the present study, we used a three-item measure of life satisfaction including both evaluative and hedonic aspects of well-being. A single-item measure of life satisfaction has previously been used (134, 135), including in our NORDOC studies (19, 20). Another instrument, the Physician Wellness Inventory, which includes the items high career purpose, lack of distress and high cognitive flexibility, was developed for use with family physicians but has not, to our knowledge, been used with those in other medical specialties (17). It is important to note that the terms ‘physician satisfaction’ and ‘professional satisfaction’ are commonly used in the USA to refer to what Europeans call ‘job satisfaction’ (136). Job satisfaction and life satisfaction are unique constructs that are closely related to life satisfaction, as noted by Heller et al. (137). The correlation between job satisfaction and life satisfaction has been reported as $r = 0.44$, indicating that they are not strongly associated (137, 138) and therefore represent distinct concepts.

Despite studies showing more work-related and emotional distress among doctors compared with the general population, most evidence comes from studies carried out during the doctor’s initial postgraduate years and/or internship/residency (43, 134, 139-141). Little is known about positive psychological outcomes, life satisfaction and overall well-being later in a doctor’s career, or among more established doctors. We expect that studies such as this thesis will capture the longer-term effects of doctors who experience unfavourable working
conditions. Furthermore, assessment of the relative impact of work-related factors on life satisfaction has not controlled for possible confounders, such as individual factors (personality, drinking behaviour, lifestyle) and stress outside of work (negative life events). These were also investigated in the present study.

1.11 Factors associated with life satisfaction

To our knowledge, few prospective longitudinal studies have addressed doctors’ work and life satisfaction (20, 142-144), and most of these have been cross-sectional (145-149). Prospective and longitudinal studies are therefore needed. We expect that the work-related factors described below may be related to doctors’ life satisfaction beyond their initial postgraduate years.

Work-related factors

We assessed the same work-related factors (perceived job stress factors, perceived job demands, and autonomy), as described above, in relation to hazardous drinking. In addition, we studied number of hours at work per week and number of hours asleep when on call, because doctors often work more than full-time (150) and sleep deprivation (151, 152) can negatively affect both their well-being and the quality of care they provide their patients (127, 153). Both factors are potential sources of stress for doctors, while reduced working hours and improved sleep among young doctors are associated with fewer medical errors and fewer failures (154, 155). In a US survey, working ≤ 60 hours/week had a positive effect on doctors’ well-being (120). However, in Norway, as in the other Scandinavian countries, working hours are tightly regulated so that very few doctors work > 60 hours/week. Previous studies have shown that Norwegian doctors are more satisfied with their working hours than are doctors in Germany, who work longer hours than their Norwegian colleagues (150, 156, 157).

Another important work-related factor is colleague support, which may buffer the negative effects of work demands (158). NORDOC researchers previously found that colleague support is protective against work–home interference stress (114). We have not yet examined whether colleague support may have a direct effect on their well-being beyond the initial postgraduate years.

Individual predictors and/or factors associated with life satisfaction
We controlled for several possible non-work confounders, which may serve as mediators and/or moderators of a possible relationship between work-related factors and life satisfaction.

**Age.** A cross-sectional US study has shown that older doctors are more satisfied with their lives than are younger doctors (148), which is consistent with a recent Irish study (159).

**Gender.** A prospective Swiss study of young doctors showed lower life satisfaction among both genders compared with a representative sample of the general population (143). Compared with male doctors, female doctors were more satisfied with their lives with regard to income (measured using a domain-specific life satisfaction scale). In a cross-sectional study of American, Canadian, and Mexican oncologists (120), women oncologists were less likely than their male colleagues to report a high degree of overall well-being (31 vs. 53%, respectively; P = 0.02). This aspect has not been studied in a more egalitarian Scandinavian society.

**Having children.** A review of general population studies in Europe and the USA showed that parents tend to be less satisfied with their lives compared with childless individuals (160).

**Being married/cohabiting and perceived social support.** Being married or cohabiting is a form of structural support. One cross-sectional study noted that married physicians were more satisfied than unmarried physicians (161). Consistent with this, a previous NORDOC report showed that perceived social support accounted for significant variance in life satisfaction (20). We therefore expect this effect to be important in the current larger and longer follow-up study of Norwegian doctors.

**Physical activity.** A growing body of literature shows that physical activity is associated with well-being among both the general population (162) and doctors (163). A national survey of both Australian and New Zealand junior doctors showed that about 27% engaged in exercise $\geq$ 4 hours/week, whereas 15% reported no exercise during the preceding month (164). The Norwegian Health Directorate recommends 150 min of moderate-intensity exercise/week or 75 min of high-intensity exercise/week (3). NORDOC researchers found that physical
training significantly predicted life satisfaction, even after controlling for anxiety and depressive symptoms (20).

**Religious activity.** Religious activity may be a coping strategy against illness and life stress (165). A cross-sectional study showed that religious activity is positively linked to life satisfaction among doctors (166). Our group found that religious activity predicted high life satisfaction during the first postgraduate year among doctors, although this became non-significant when controlling for social support (20). Nevertheless, further studies examining the impact of religious activity among more career-established doctors are required.

**Use of alcohol to cope with tension and hazardous drinking.** Both of these variables were described above in the section 1.9, and they were also assessed to determine their link to doctors’ life satisfaction. Drinking plays an important part in daily social contexts and may have a negative impact on overall life satisfaction. However, doctors also enjoy drinking and engage in this for recreational purposes. Furthermore, abstainers have slightly higher mortality rates compared with those who drink moderately, whereas heavy drinkers have the highest mortality rates (167, 168). We therefore included both drinking to cope with tension and hazardous drinking in our statistical models, to determine whether either or both are independently related to doctors’ life satisfaction.

**Personality traits.** Within the general population, the personality traits of extraversion and conscientiousness have been positively linked to life satisfaction, whereas high neuroticism is negatively linked to life satisfaction (169). In the same survey, personality traits and perceived social support accounted for about 37% of variance in life satisfaction. To our knowledge, only two prospective longitudinal studies have looked at personality and positive psychology (e.g., life satisfaction) among doctors. The first was a 12-year cohort study of doctors in the UK, which showed that lower neuroticism predicted better overall satisfaction with medicine five years post-graduation (170). The second was a prospective cohort study by the NORDOC group, showing that low levels of neuroticism predicted better life satisfaction among Norwegian doctors nine years post-graduation (20). The personality trait of neuroticism appears to be important with respect to positive psychological outcomes (e.g., life satisfaction) among both doctors and the general population; this might be due to a common genetic basis (171, 172). Nevertheless, more longitudinal studies are needed to
assess the effects of personality traits, beyond doctors’ training or initial postgraduate years. Personality measures have rarely been included in epidemiological studies of doctors, presumably because many personality inventories are quite long. One strength of the present study is that personality was measured during medical school, allowing us to assess its predictive validity of life satisfaction within this population.

**Life events.** We controlled for this non-work-related variable, which is associated with life satisfaction among Norwegian doctors (20). We anticipated that negative life events (e.g., divorce, financial problems, illness) would increase over the years, and thus expected these to negatively impact doctors’ quality of life.
2. **AIM OF THE THESIS**

The aim of this thesis was to determine the prevalence of, and individual- and work-related factors contributing to, hazardous drinking and life satisfaction using a longitudinal nationwide sample of Norwegian medical doctors. The sample was followed for 15 years after leaving medical school. Ultimately, the goal of this research programme is to improve medical education about individual- and work-related factors associated with hazardous drinking and life satisfaction among doctors. In addition, this work may help identify organizational- and work-related factors that can be targeted for interventions in Norwegian doctors’ workplaces.

2.1 **Research questions**

1) What are the prevalence rates and stability of hazardous drinking at the end of medical school and post-graduation, and which individual factors measured during medical school predict later hazardous drinking? (Papers I and II)

2) Are doctors’ work stress, mental distress and/or life stress independent factors related to their hazardous drinking? (Paper II)

3) What is the level of life satisfaction among Norwegian doctors, which work-related factors are independently associated with life satisfaction, and which individual- and work-related factors predict a decline in life satisfaction over the first 15 years of a doctor’s career? (Paper III)

4) Are doctors’ drinking behaviours (i.e., hazardous drinking and drinking to cope) independently linked to their life satisfaction? (Paper III)

5) Are there any gender or cohort differences in variables predicting doctors’ hazardous drinking and/or life satisfaction? (Papers II and III)
3. MATERIALS AND METHODS

3.1 Historical background of the NORDOC project: The Longitudinal Study of Norwegian Medical Students and Doctors

During the early 1990s, there was significant media coverage of physician stress and suicide. The Norwegian Medical Association (NMA) was concerned about this, and invited Professor Bengt Arnetz from Sweden to give a talk about his research on Swedish doctors’ working conditions (173-175). The NMA decided that similar studies should be undertaken in Norway, and that a research programme should be initiated. ‘Legekårsundersøkelsen’ (Norwegian Physician Health Study) was established and funded by the Nursing and Pension Scheme for Doctors (SOP). In the spring of 1992, Olaf G. Aasland began as the programme director, assisted by Erik Falkum, psychiatrist and senior researcher, and Wenche Hvitmyhr, project manager. Various academic professionals were invited to submit additional project proposals. Professors Per Vaglum and Øivind Ekeberg at the Department of Behavioural Sciences in Medicine, Medical Faculty, University of Oslo, had completed some research on medical students and proposed a nationwide longitudinal study of medical students and their subsequent careers. After obtaining permission from the Norwegian Data Inspectorate, a large-scale, comprehensive follow-up study commenced. That study included medical students at all four Norwegian universities who either started their studies in 1993 (the Medical Student Cohort) or finished their studies during 1993 or 1994 (the Young Doctor Cohort). To increase participation, Aasland, along with medical student and politician Nina Tangnæs Grønvold (later to become State Secretary for the Minister of Education and Research) visited all universities in 1993.

Data collection for that study was a success, and the first researcher to use the data was sociologist Jannecke Wiers-Jenssen, for her master’s thesis ‘Recruitment to Medical Study. A Study of a Class of Medical Students with Emphasis on Social Backgrounds, Collection of Study Points, Study Motives and Ambition Level’ (176). In 1995, psychiatrist Reidar Tyssen took over primary professional responsibility for the study follow-ups and continued both cohorts under the supervision of Professors Vaglum and Ekeberg. Tyssen became the first to use the novel follow-up data, which was later renamed NORDOC in his PhD thesis ‘Mental Health Problems among Medical Students and Young Physicians: A Nationwide and Longitudinal Study’ (50).
Thus, two institutions were responsible for this research on the health and working conditions of Norwegian doctors: the Institute for Studies of the Medical Profession (LEFO), which was linked to the Medical Association’s Secretariat and funded by the NMA and SOP, and NORDOC, based in the Department of Behavioural Sciences in Medicine at the University of Oslo, supported by a small annual grant from LEFO. Over the years, various initiatives have been undertaken to maintain the high response rate originally achieved in NORDOC. The consent form administered prior to the survey may itself create a sense of commitment. Furthermore, NORDOC staff reward participation by sending a music CD to every respondent. The high response rate may also have been due to the ‘duty fulfilment’ attitude, a particularly common characteristic of medical students and doctors. Another reason for NORDOC’s high response rates may be the frequency of media reports on doctors’ health over recent decades. With regard to baseline recruitment, the importance of the visit in 1993 of Grønvold and Aasland to the four schools cannot be underestimated. Ninety-six per cent of registered Norwegian physicians are members of the NMA, and the agency’s approval and acknowledgement in the study cover letter may have been another reason for loyal participation in NORDOC (Olaf G. Aasland, personal communication).

3.2 Study design

NORDOC is a nationwide longitudinal prospective study covering a 20-year period for each of two cohorts: the Medical Student Cohort (N = 421) and the Young Doctor Cohort (N = 631). Baseline measurements of these cohorts were collected approximately six career years apart, at the beginning and end of medical school, respectively. Both cohorts were approached via postal surveys (30–40 pages of questionnaires) on six different occasions during the 20-year span (1993 to 2014), with later follow-ups at roughly five-year intervals. In this thesis study, we used NORDOC data from both cohorts at each of four time points during the 15-year follow-up (Fig. 2).

3.3 Sample

The NORDOC sample consists of both the Medical Student Cohort (N = 421) and the Young Doctor Cohort (N = 631). The Medical Student Cohort was all students at each of the four Norwegian universities who entered medical school in 1993; the Young Doctor Cohort was all medical students who graduated from any of the four Norwegian universities in 1993 or 1994. The last NORDOC survey follow-up was in 2014, at which time most of the doctors had completed postgraduate training (note that it takes approximately nine years after
physician authorization to complete training for a first specialty; 85% of specialists complete training within 5–15 years after authorization) (177, 178). For the present study, we merged data from these cohorts for the first time, to increase sample size and allow exploration of cohort effects on work-related predictors.

With these cohorts merged (N = 1052) at each of four similar career time points, our subsamples and response rates were as follows: at the final year of medical school (T1; data collected during 1993/1994 and 1999, respectively; N = 892/1052, 85% of the eligible sample); four years after graduation (T2; data collected in 1998 and 2003, respectively; N = 780/1052, 74% of the eligible sample); 10 years after graduation (T3; data collected in 2003 and 2008, respectively; N = 708/1052, 67% of the eligible sample); and 15 years after graduation (T4; data collected in 2008 and 2014, respectively; N = 598/1052, 57% of the eligible sample). Overall, 90% (947/1052) of the eligible sample responded at least once and 42% (450/1052) responded at all four time points; 59% (264) of this sample were women and 41% (186) men. At T1, their average age was 28 (standard deviation [SD] = 2.8) years and gender distribution was 56% (n = 502) women and 44% (n = 390) men. Figure 2 illustrates the sample and study design details.
T1 = 892/1052 (85% Rr)
T2 = 780/1052 (74% Rr)
T3 = 708/1052 (67% Rr)
T4 = 598/1052 (57% Rr)
90% responded at least once; 42% responded at all four time points
Rr = Response rate

Fig. 2. Sample and study design (N = 1052).

3.4 How representative of Norwegian medical doctors is our sample?
We compared NORDOC study participants with doctors who joined the NMA in 2014, at a similar mean age, to determine how representative our sample is of Norwegian doctors. In 2014, there were approximately 27,900 doctors in Norway (including approximately 1,000 non-NMA-member doctors, i.e., 96% are NMA members). In Norway, 98% of doctors (including specialists) work in the public health system or are funded by the government and social security system; only 2% of doctors (i.e., general practitioners and specialists) are in private practice without public support (self-pay patient practice).

NORDOC is representative of medical students from the four Norwegian universities, but not of currently practising doctors due to considerable immigration, including among doctors, over the past two decades. In our sample of pooled NORDOC cohorts, there was obvious skewedness with respect to age; our sample’s average age was 43 (SD = 2.8) years at T4, while the 2014 NMA sample was 44 (SD = 11.5) years. The median age for both the combined NORDOC cohorts and the NMA samples was 42 years. About 93% of the
NORDOC doctors at T4 were 40–49 years old, 59% of whom were women and 41% men, whereas there was a 50:50 ratio in the NMA sample.

With respect to the numbers of doctors in different specialties, we found significant differences between the NORDOC and NMA samples in the specialties: general practitioners (including family practice) (22% vs. 26%, respectively, $P = 0.032$), internal medicine specialties (27.5% vs. 21%, respectively, $P < 0.001$), and paediatrics (6.7% vs. 4.0%, respectively, $P = 0.005$). No significant differences were found with respect to surgical specialties (including anaesthesiology), psychiatry (including adolescent and child psychiatry), obstetrics/gynaecology or other specialties (e.g., laboratory medicine, immunology, etc.) (Table 1). Further comparisons showed no significant differences in the numbers of hospital-based and non-hospital-based doctors between NORDOC at T4 (63% vs. 37%, respectively) and the NMA sample (61% vs. 39%, respectively). Non-hospital-based doctors were mainly general practitioners, but also included municipal doctors with and without leadership responsibilities, private practice doctors, company doctors, full-time researchers and doctors working in administration. We can conclude that NORDOC sample doctors working in surgical specialties, obstetrics/gynaecology and other specialties (e.g., laboratory medicine, immunology, etc.) are highly representative of the population of Norwegian doctors. However, compared with the overall population, the NORDOC sample includes relatively fewer general practitioners and more doctors in internal medicine and paediatrics.
### Table 1. NORDOC sample (T4) versus NMA population (2014) according to generalist or specialty status

<table>
<thead>
<tr>
<th>Specialty</th>
<th>NORDOC (%)</th>
<th>NMA (%)</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPs (family/general practice)</td>
<td>22.0%, N = 105/480</td>
<td>26.0%, N = 6066/23,039</td>
<td>0.032</td>
</tr>
<tr>
<td>Internal medicine specialties</td>
<td>27.5%, N = 132/480</td>
<td>21.0%, N = 4811/23,039</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Surgical (including anaesthesiology)</td>
<td>18.5%, N = 89/480</td>
<td>21.4%, N = 4937/23,039</td>
<td>0.141</td>
</tr>
<tr>
<td>Psychiatry (including adolescent and child psychiatry)</td>
<td>10.4%, N = 50/480</td>
<td>11.0%, N = 2554/23,039</td>
<td>0.697</td>
</tr>
<tr>
<td>Obstetrics/gynaecology</td>
<td>4.4%, N = 21/480</td>
<td>4.5%, N = 1039/23,039</td>
<td>0.949</td>
</tr>
<tr>
<td>Paediatrics</td>
<td>6.7%, N = 32/480</td>
<td>4.0%, N = 925/23,039</td>
<td>0.005</td>
</tr>
<tr>
<td>Others (laboratory medicine, immunology, etc.)</td>
<td>10.6%, N = 51/480</td>
<td>11.7%, N = 2707/23,039</td>
<td>0.997</td>
</tr>
</tbody>
</table>

NORDOC = The Longitudinal Study of Norwegian Medical Students and Doctors  
NMA = Norwegian Medical Association  
*Chi-squared test

#### 3.5 Dependent variables

##### 3.5.1 Hazardous drinking (single-item measurement)

In Paper I, *hazardous drinking* (binge drinking, drinking to intoxication and/or problematic drinking) was defined as drinking a total 60 g of ethanol (approximately five alcoholic units) or more in one session, at least 2–3 times/month in the past year. The quantity of alcoholic beverages was specified (i.e., five half-bottles [330 ml] of beer [4.5% of ethanol], one bottle of red or white wine, half a bottle of dessert wine, ¼ bottle of brandy/whisky). Responses were: none = 0; 1–4 times/year = 1; 5–10 times/year = 2; 1 time/month = 3; 2–3 times/month = 4; once a week = 5; 2–4 times/week = 6; daily or almost daily = 7. This single-item hazardous drinking measure is a modified version of the third item (9) on the original 10-item version of the AUDIT (179). Unfortunately, we did not include the entire AUDIT in our questionnaire battery (i.e., at T1, T2 or T3), although this single item has been determined to be the best AUDIT screening item for detecting alcohol problems among students and men in general medical settings (180-182).

Before performing statistical analyses, we dichotomized hazardous drinking as ≤ 1 time/month = 0 or ≥ 2–3 times/month = 1, because this cut-off corresponds to the frequency
of hazardous drinking (every fortnight or more often) among US college students linked to negative health consequences (12). We also validated this cut-off in our previous research (9-11). The same cut-off was used for men and women. However, compared with men, women need less alcohol to become intoxicated, based on their lower body fluids and gastric alcohol dehydrogenase, the enzyme for alcohol metabolism (183). For this reason, and to avoid under-estimation of hazardous drinking among women, we performed an internal validation with our cut-off.

Validation of the single-item measure of hazardous drinking

We validated our single-item hazardous drinking measure using medication self-prescribing data (184). We compared use of medications for three common health problems associated with alcohol abuse and dependency: hypertension (13), gastritis (14), and sleep problems (15). At T2, hazardous drinkers reported significantly higher use of antacids compared with non-hazardous drinkers (23.4% vs. 11.4%, respectively, P = 0.006), higher use of antihypertensive drugs (3.9% vs. 0.3%, respectively, P = 0.008) and had a trend towards higher use of hypnotics (13.0% vs. 6.9%, respectively, P = 0.06). At T3, hazardous drinkers reported significantly higher use of antacids (30.2% vs. 15.0%, respectively, P = 0.01) and hypnotics (20.8% vs. 9.0%, respectively, P = 0.01), but not of antihypertensive drugs (0.0% vs. 1.2%, respectively, P = 1.00). This single-item hazardous drinking variable measured at T1 was also used to control for baseline drinking in Paper I.

To examine the validity of the single-item measure of hazardous drinking (i.e., drinking a total of 60 g of ethanol ≥ 2–3 times/month within the past year) among women, we compared responses on this item to responses on the item: having four drinks on the same occasion (measured only in the Medical Student Cohort at T3) and found that these items were highly correlated (r = 0.83, P < 0.001).

We also assessed the internal validity of the hazardous drinking item, based on another screening instrument, the CAGE (Cut down drinking, Annoyed by criticism, Guilty feelings, and Eye opener), which consists of four simple questions and binary response options (Yes/No) (185). At T3, the CAGE was measured in the Medical Student Cohort (N = 318/500) and our sample size was therefore reduced. Hazardous drinking was defined as a score ≥ 2 on the CAGE and was correlated with our single-item measure of hazardous drinking (r = 0.43, P < 0.001), indicating fair validity.
3.5.2 Hazardous drinking (nine-item version)

In Paper II, hazardous drinking was measured using a modified nine-item version of the AUDIT (103). The original AUDIT comprised 10 items across three domains: alcohol consumption (items 1–3), alcohol dependence (items 4–6), and alcohol-related consequences (items 7–10), with each item scored on a 0–4-point scale, giving a maximum score of 40 (179). Unfortunately, as noted, we did not collect the full AUDIT but instead used a modified nine-item version (103). The third AUDIT item, ‘How often do you have six or more drinks on one occasion?’ was not included in the questionnaire battery at T1, T2 or T3 because this type of drinking behaviour was unfamiliar to young Norwegians during the 1990s, when drinking liquor and spirits was uncommon compared with drinking beer or wine (9). In the Medical Student Cohort, we administered the 10-item version at T4, which was strongly correlated with the nine-item version ($r = 0.98$, $P < 0.001$). As noted, in the literature, a different cut-off has been used with the 10-item AUDIT to define hazardous drinking (30, 186). In the nine-item version, we used a cut-off score of $\geq 6$ for men and $\geq 5$ for women, which have been barely validated (61, 103).

Validation of our AUDIT-9 cut-off scores for hazardous drinking

Wurst et al. (63) questioned the validity of the typical AUDIT cut-off scores defining hazardous and harmful drinking among doctors, suggesting that higher thresholds are more valid. Because we lacked validity information for the cut-off at $\geq 6$ for men and $\geq 5$ for women on the nine-item version, we performed an internal validation in our sample using these cut-off scores with respect to several common alcohol-related medical conditions that we assessed using self-report: use of antihypertensive, antacid and hypnotic medications. At T2, T3 and T4, hazardous drinkers (cut-off $\geq 6$ for men and $\geq 5$ for women on the nine-item version) reported significantly higher use of antacids compared with non-hazardous drinkers (T2: 28.4% vs. 10.2%, $P < 0.001$; T3: 25.3% vs. 14.9%, $P = 0.021$; T4: 28.7% vs. 16.4%, $P = 0.010$) and hypnotic drugs (T2: 14.7% vs. 6.4%, $P = 0.007$; T3: 17.6% vs. 8.7%, $P = 0.013$; T4: 21.6% vs. 8.6%, $P = 0.001$). We identified a significantly higher use of antihypertensive drugs among hazardous drinkers compared with non-hazardous drinkers at T2 (2.9% vs. 0.3%, respectively, $P = 0.019$), but not at T3 or T4. Since both our single-item and the nine-item AUDIT versions are associated with drinking in our sample, we assume these to be highly valid.
3.5.3 Life satisfaction
In Paper III, life satisfaction was measured at T2, T3 and T4 using three items. Item 1 asked, ‘When you think about your life today, would you say that you are by and large satisfied with life, or are you mostly dissatisfied?’ with response options ranging from 1 (‘extremely dissatisfied’) to 7 (‘very satisfied’). This and similar items have been used and validated previously (20, 134, 135). Item 2 asked, ‘To what extent are you satisfied with your daily life situation?’ with response options ranging from 1 (‘very dissatisfied’) to 4 (‘very satisfied’). Item 3 asked, ‘Would you describe yourself mostly as…?’ with response alternatives ranging from 1 (‘not at all happy’) to 5 (‘very happy’). Our factor analysis of these three items at T2 yielded an unequivocal unidimensional solution, so that the factor score coefficients were multiplied by their respective constituent items (raw scores) and summed to create a simple additive scale. The factor score coefficients from T2 were also used at T3 and T4 (i.e., without standardizing the raw scores before multiplying and summing). This procedure allowed us to assess individual-level changes and thus use a mixed-model repeated-measures (i.e., longitudinal) approach. Cronbach’s alphas for this scale were T2 = 0.83, T3 = 0.83 and T4 = 0.82, indicating very good reliability within this relatively brief, three-item instrument.

3.6 Independent variables

3.6.1 Socio-demographic variables (age, gender, married/cohabitating)
Age was measured as a continuous variable in years. The average age of the sample at T1 was 28 (SD = 2.8) years.

Gender was coded as female = 0 and male = 1 in Papers I and II, whereas in Paper III it was coded as male = 0 and female = 1.

Married/cohabitant status was measured at T2, T3 and T4 with one question regarding current marital status and six response alternatives: unmarried, separated, widow/widower, divorced = 0 and married and cohabiting = 1 (20, 116). This variable measures structural social support, such as having a stable partner, as well as the negative effect of being divorced or separated (coded 0). This variable was included only in Paper III but has been validated in previous research (20).
3.7 Other predictor variables

3.7.1 Having children or not

*Having children or not* was measured with the question, ‘How many children do you have?’ and response categories none = 0; one child = 1; two children = 2; and three or more children = 3. We dichotomized this variable as no children = 0 or one or more children = 1 in Papers I and III; in Paper II we coded this variable as one or more children = 0 or no children = 1. This variable has been previously validated by our group (9).

3.7.2 Religious activity

*Religious activity* was measured with the question, ‘Do you engage in any type of religious activity?’ and response options No = 1; Yes, I read religious literature almost daily = 2; Yes, I sometimes read religious literature = 3; Yes, I frequently attend services and religious meetings = 4; Yes, I sometimes attend services and religious meetings = 5; and Yes, I actively participate in other types of religious activities = 6. We dichotomized this variable as no religious activity = 1 or participation in any type of religious activity = 0 to test the effects of no religious activity on hazardous drinking (Papers I and II). In Paper III, this dichotomous variable was coded inversely to determine whether any type of religious activity had an effect on life satisfaction. The reliability of this single-item variable may be limited, although it was previously validated (9).

3.7.3 Perceived social support

*Perceived social support* was measured at T2, T3 and T4 using five questions about the amount of appreciation received from close friends, the presence of warm and caring confidants, the degree of affiliation with groups such as one’s neighbourhood, political organization, or church, and the support anticipated if one should fall ill. All items were scored with one of five response categories and the total score could range from 1 to 5, with higher scores indicating higher levels of experienced support. The total score of all five items was used to assess the level of experienced support in the present study.

Regarding reliability, Cronbach’s alphas for this scale were 0.91, 0.71 and 0.72 at T2, T3 and T4, respectively. We cannot explain the reduction in reliability from T2 to subsequent years, but it was satisfactory at each time point. This variable was included only in Paper III but has been validated as an important predictor of life satisfaction (20).
3.7.4 Physical activity

Physical activity was measured at T2, T3 and T4 with the question, ‘Do you usually do physical exercise or workout, such as jogging, long hikes, cross-country skiing, aerobics, cycling, swimming, football, tennis, etc.?’ with five response categories: no = 0; less than once weekly = 1; yes, 1–2 days per week = 2; 3–4 days per week = 3; and yes, 5–7 days per week = 4. This question has been previously validated (20) and was included only in Paper III.

3.7.5 History of parental alcohol problems

History of parental alcohol problems (i.e., having parents with alcohol problems) was measured at T1 by asking, ‘Have your parents had alcohol problems to such an extent that it has had negative consequences on their health, family life or work?’ with response categories: no = 0; yes, my mother = 1; yes, my father = 2; and yes, both parents = 3. We dichotomized this variable as neither parent = 0 or one or both parents = 1 to determine whether parental drinking (mother, father, both) predicted later drinking in Papers I and II.

3.7.6 Cohort

Cohort was coded as Medical Student Cohort = 0 or Young Doctor Cohort = 1 in Papers I and II; while in Paper III they were coded the opposite. In other words, in Papers I and II, the cohort variable is positive for the older cohort and the inverse for Paper III.

3.7.7 Personality traits

Personality traits were measured using the 36-item version of Torgersen’s Basic Character Inventory (BCI), which is Torgersen’s (187) modification of the original questionnaire constructed by Lazare et al. (188). This instrument assesses four personality dimensions: vulnerability, intensity, control and reality weakness, which are the terms used in Paper I. Each BCI dimension is assessed using nine questions, each with dichotomous response options (‘agree’/‘do not agree’), and a total ranging from 0 (low) to 9 (high). The first three BCI dimensions resemble Eysenck’s Giant Three (189): neuroticism, extraversion and conscientiousness, respectively; these terms are used in Papers II and III. The BCI vulnerability dimension resembles the classic neuroticism scale, with items related to low self-esteem or self-criticism (e.g., ‘I am very touchy about criticism’, ‘It often seems that others do things much better than me’). The BCI intensity dimension measures extraversion/introversion (e.g., ‘Many people consider me a lively person’, ‘It is rare I get
particularly excited’, ‘When I am with other people, I like to be in the background’ [the last two examples are reverse scored]. The BCI control dimension measures the degree of obsessiveness or conscientiousness (e.g., ‘Everything I do must be precise and accurate’, ‘My accuracy and sense of order have occasionally annoyed other people’). Reality weakness on the BCI is an original, deviant trait related to perceptions and ideations on the borderline between reality and weakness; this dimension also measures chronic illusions, paranoid traits and traits related to severe personality disorders (190, 191). Examples of reality weakness items are, ‘I experience myself as being totally different at different points in time’, ‘I feel lonely most of the time’, and ‘Sometimes I feel I am not myself’.

Personality traits were measured in a randomly selected half of the Young Doctor Cohort sample at T1 and one year later in the remaining half. This was done for practical reasons and because we needed to reduce the volume of the questionnaire battery. No significant differences were found between these subsamples on \( t \) tests. Stability of the vulnerability dimension across this year was tested with the ‘Low general self-esteem’ items from the BCI vulnerability dimension; the test–retest correlation was high \( (r = 0.7) \) (192).

In the Medical Student Cohort, personality traits were measured in the whole sample upon entrance to medical school. The vulnerability (neuroticism) and intensity (extraversion/introversion) dimensions had the highest reliability (Cronbach’s alphas > 0.70), whereas control (conscientiousness) (Cronbach’s alpha = 0.63) and reality weakness (Cronbach’s alpha = 0.62) were relatively lower (50), showing poorer reliability. Additional analyses provided evidence that these lower alphas may be due to their dichotomous response format (‘agree’/‘not agree’), which could have been improved by using a five- or seven-point Likert scale (Torbjørn Moum, personal communication). The BCI has been used in many previous studies (190, 193, 194) and the short version has been validated in studies of Norwegian medical students and doctors (116, 195-198). Although the BCI has not been validated for use internationally, three of its dimensions (i.e., subscales) closely resemble personality inventories that have been used internationally by well-established researchers including Eysenck, Cloninger, and Costa and McCrae. The most frequently used and best-validated personality measure is Costa and McCrae’s ‘Big Five’, which captures the two additional dimensions agreeableness and openness, in addition to neuroticism, extraversion and conscientiousness (199).
3.7.8 Use of alcohol to cope with tension

Use of alcohol to cope with tension was measured by the question, ‘When you feel worried, tense, or nervous, do you ever drink alcoholic beverages to help you handle things?’, with four response options: never = 0; seldom = 1; occasionally = 2; and often = 3. This question was previously used and validated with a large US sample (200) and in mental health surveys of students and doctors in Norway (9, 10, 103). The reliability of this variable may be somewhat low because it was measured by a single-item. This type of behaviour in Norwegian culture is stigmatized and the only answer among responders would be ‘never’. On this basis, and to determine whether such a drinking strategy was ever used, we dichotomized this variable as never = 0 or 1 = yes (any frequency). Another reason for dichotomizing was that the responses were skewed despite two attempts at log transformation. Nevertheless, our use of a single item reduces its reliability and increases the risk of type II error.

3.7.9 Life events

Life events during the past year were measured at T2, T3 and T4 with 13 items:

   1) Serious disease/accident/hospital admission
   2) Divorce/separation/broken relationship
   3) Got married/started living with a cohabitant
   4) Had children
   5) Death of family member/close friend
   6) Other difficulties in the immediate family
   7) Serious financial problem
   8) Serious problems with residence/dwelling
   9) Partner unemployed/granted leave
  10) You, or someone in the immediate family, involved in serious law violation
  11) Problem with partner
  12) Moved away from parents
  13) Other serious event (self-specified)

The life events variable was assumed to be either positive (e.g., having a child) or negative (e.g., having a serious illness/accident/hospital admission) to the respondent’s well-being. All items were coded 0 or 1, and their sum (both positive and negative) was used in analyses; we also performed post hoc analyses to identify which events were significant.
predictors. This measure has been used in previous studies of the two NORDOC cohorts (45, 116, 201) and therefore we consider it well validated. In Paper II, a weighted sum of significant life events was also assessed as a predictor of hazardous drinking; this use increased reliability of these data and validity of the measure. Nevertheless, the total score could be considered a mixture of ‘apples and bananas’ (i.e., negative and positive events). Variable life events were first studied by Holmes and Rahe in 1967 (202).

3.7.10 Mental distress

Mental distress during the past two weeks was measured at T2, T3 and T4 using the five-item Symptom Checklist (SCL-5), a shorter version of the SCL-25 which is based on a factor analysis by Tambs and Moum (203). The SCL-5 and SCL-25 are significantly correlated ($r = 0.92$) and had strong reliability (Cronbach’s alpha = 0.85). Three SCL-5 items relate to anxiety (items 1, 2 and 5) and two to depression (items 3 and 4). The SCL-5 also includes an item about how much the respondent was bothered by each of the five symptoms during the past two weeks. The symptoms are: 1) ‘Feeling fearful’; 2) ‘Nervousness or shakiness inside’; 3) ‘Feeling hopeless about the future’; 4) ‘Feeling blue’; and 5) ‘Worrying too much about things’, each of which is measured on a five-point scale where 1 = ‘not at all’ and 5 = ‘very much’. The mean item score was used to assess the level of mental distress. Cronbach’s alphas for this scale were 0.87 at T2, 0.87 at T3 and 0.85 at T4, indicating good reliability. This measure of mental distress has been used and validated in several studies of the general population (203-205) and of Norwegian students and doctors (9, 20, 206, 207). It is important to note that this mental distress variable measures psychological distress (i.e., anxiety and depressive symptoms) rather than severe depression or anxiety disorders. In Paper II, we attempted to separate the anxiety and depression dimensions, but found the subscales to be very highly correlated ($r = 0.68$, $P < 0.001$). The SCL-5 we used here differs from alternative four-item versions and, when dichotomized, requires a different cut-off level. This version was chosen because it is similar to a version used for the Norwegian Physician Health Survey and the Villa Sana study, and has been well validated (206).

3.7.11 Work-related variables

Perceived job stress, used in Papers II and III, was measured at T2, T3 and T4 with a modified version of the Cooper Job Stress Questionnaire. This self-report questionnaire was originally developed by Cooper et al. (208) and later modified into a 32-item scale by Tyssen et al., the details of which have been previously described (116). Each item was measured on
a five-point scale (1 = ‘no stress’ to 5 = ‘a lot of stress’). Our group’s previous principal component analysis of these data at T2 (114) revealed four dimensions: 1) Emotional pressure (Cronbach’s alpha = 0.83; eight items, e.g., ‘I am stressed by dealing with suffering patients’, ‘having 24 hour responsibility for patients’ lives’, ‘dealing with problem patients’); 2) Time pressure (Cronbach’s alpha = 0.71; six items, e.g., ‘I am stressed by interruptions and fuss at work’, ‘medical records and paperwork’, ‘night calls’, ‘working environment’, ‘admissions’); 3) Fear of complaints/criticism (Cronbach’s alpha = 0.75; seven items, e.g., ‘I am stressed by worrying about patients’ complaints’, ‘conflicts with colleagues or other staff’, ‘expectations that the physicians should deal with non-medical problems’); and 4) Work–home interference (Cronbach’s alpha= 0.86; three items, e.g., ‘demands of your job on family life’, ‘balancing oneself between work and private life’, ‘demands of your job on social life’). The mean item score for each dimension was used as a measure of perceived job stress level. Regarding reliability, the internal consistency (Cronbach’s alphas of the dimensions, 0.71–0.86) was good. This modified perceived job stress questionnaire has been validated in several NORDOC studies (114, 116, 198, 207).

In Papers II and III, two self-reported and presumably more ‘objective’ workload measures were collected at T2, T3 and T4: average number of working hours per week and average number of hours asleep when on call. Both measures are single item and therefore their reliability values may be limited.

In Papers II and III, psychosocial factors including perceived job demands and autonomy were measured at T2, T3 and T4 with 10 statements, the details of which have been described previously (209). Our principal component analysis of these 10 items at baseline and T2 identified two factors: perceived job demands (eight items, Cronbach’s alpha = 0.82) and autonomy (two items, Cronbach’s alpha = 0.78). The summed score of each factor was used in analyses. Examples of questions or statements assessing perceived job demands are, ‘Do you sometimes have so much to do that your work situation becomes hurried and fussy and, if so, how often?’; ‘You work under unacceptable pressure’, and ‘You have so many tasks that they negatively affect your ability to work effectively’. The two autonomy items are, ‘To what extent can you decide your work pace?’ and ‘To what extent can you decide or plan the order of your tasks during the day?’ The reliability of perceived job demands and autonomy according to their Cronbach’s alpha values was quite good. As noted, these measures resemble Karasek’s demand–control variables in the QPS Nordic questionnaire (118). These demand–control variables were recently implemented at T4 (in the Medical
Student Cohort); there was a significant correlation between our variables and those of Karasek with respect to job demands ($r = 0.37, P < 0.001$) and control or autonomy at work ($r = 0.53, P < 0.001$). However, this correlation was not high ($< 0.7$), indicating that these factors resemble, but are not identical to, the demand–control variables.

In Paper III, colleague support was measured at T2, T3 and T4 using two items: ‘To what degree are you taken care of by your colleagues?’ and ‘To what degree do you enjoy working with your colleagues?’ Each item was measured on a seven-point scale (from 1 = ‘not at all’ to 7 = ‘to a very high degree’). The summed score was used as a measure of level of colleague support. Cronbach’s alphas for this scale were T2 = 0.80, T3 = 0.83 and T4 = 0.85, indicating very good reliability for all three items. This variable has been validated in previous studies (114, 210). Table 2 shows the independent and dependent variables used in Papers I–III. Table 3 describes all independent and dependent variables.
Table 2. Independent and dependent variables included in Papers I–III

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>Paper I</th>
<th>Paper II</th>
<th>Paper III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Gender</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Having children (≥ 1 child) or no children</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Married/cohabiting</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Perceived social support</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Physical activity</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Religious activity or no religious activity</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Having parent(s) with alcohol problems</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Use of alcohol to cope with tension</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Life events</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Mental distress (SCL-5)</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

**Personality traits, BCI**

<table>
<thead>
<tr>
<th>Personality traits, BCI</th>
<th>Paper I</th>
<th>Paper II</th>
<th>Paper III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vulnerability (Neuroticism)</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Intensity (Extraversion/Introversion)</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Control (Conscientiousness)</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Reality weakness</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

**Contextual work-related variables**

<table>
<thead>
<tr>
<th>Contextual work-related variables</th>
<th>Paper I</th>
<th>Paper II</th>
<th>Paper III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotional pressure</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Time pressure</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Fear of complaints and criticism</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Work–home interference</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Number of hours at work/week</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Number of hours asleep when on call</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Perceived job demands</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Autonomy</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Colleague support</td>
<td></td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>

**Dependent variables (outcome variables)**

<table>
<thead>
<tr>
<th>Dependent variables (outcome variables)</th>
<th>Paper I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hazardous drinking (single item)</td>
<td>x</td>
</tr>
<tr>
<td>Hazardous drinking (nine items)</td>
<td>x</td>
</tr>
<tr>
<td>Life satisfaction (three items)</td>
<td></td>
</tr>
</tbody>
</table>

SCL-5 = Symptom Checklist-5
BCI = Basic Character Inventory
Table 3. Independent and dependent (i.e., outcome) variables measured at each of the four data collection periods

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>T1 Final year of graduation</th>
<th>T2 4 years post-graduation</th>
<th>T3 10 years post-graduation</th>
<th>T4 15 years post-graduation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (continuous)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender, male, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Having children (≥ 1 child), %</td>
<td>26\textsuperscript{a}</td>
<td>55</td>
<td>82</td>
<td>91</td>
</tr>
<tr>
<td>Married/cohabiting, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived social support</td>
<td>1.97 (1.24)</td>
<td>2.79 (0.85)</td>
<td>2.66 (0.88)</td>
<td></td>
</tr>
<tr>
<td>Physical activity</td>
<td>1.80 (1.01)</td>
<td>1.89 (1.01)</td>
<td>2.17 (1.02)</td>
<td></td>
</tr>
<tr>
<td>Religious activity, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Having parent(s) with alcohol problems, %</td>
<td>8</td>
<td>27</td>
<td>26</td>
<td>28</td>
</tr>
<tr>
<td>Use of alcohol to cope with tension, %</td>
<td>13</td>
<td>8</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Life events</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mental distress (SCL-5)</td>
<td>1.50 (0.64)</td>
<td>1.48 (0.63)</td>
<td>1.44 (0.58)</td>
<td></td>
</tr>
<tr>
<td>Personality traits, BC1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotional pressure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time pressure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fear of complaints and criticism</td>
<td>2.46 (1.002)</td>
<td>2.42 (1.002)</td>
<td>2.33 (0.99)</td>
<td></td>
</tr>
<tr>
<td>Work–home interference</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of hours at work/week</td>
<td>45.29 (9.91)</td>
<td>42.99 (8.89)</td>
<td>42.21 (10.96)</td>
<td></td>
</tr>
<tr>
<td>Number of hours asleep when on call</td>
<td>4.52 (2.79)</td>
<td>5.38 (6.36)</td>
<td>6.41 (7.14)</td>
<td></td>
</tr>
<tr>
<td>Perceived job demands</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Autonomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colleague support</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Dependent variables    |                             |                             |                             |                             |
| Hazardous drinking (single item), % | 14\textsuperscript{c} | 10                          | 8                           |                             |
| Hazardous drinking (nine items), %\textsuperscript{a} | 16                          | 14                          | 14                          | 15                          |
| Life satisfaction (three items) | 4.55 (0.82) | 4.60 (0.81)                | 4.62 (0.81)                |                             |

Data shown as percentages or means (SD) unless otherwise specified.

\textsuperscript{a} In Paper II, having no children measured at T1 was included as a predictor variable, 74% had no children.

\textsuperscript{b} In Papers I and II, having no religious activity measured at T1 was included as a predictor variable, 70% were not involved in any kind of religious activity.

\textsuperscript{c} In Paper I, hazardous drinking (single item) measured at T1 was included as a predictor variable.

\textsuperscript{d} Hazardous drinking (nine items) measured at T2, T2 and T4 was included in Paper III as an independent variable.
3.8 Statistical analyses

In Papers I and II, statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS) version 20.0; in Paper III, analyses were performed using SPSS version 23. P < 0.05 was considered statistically significant in all three papers.

**Paper I:** Paper I is a 10-year follow-up study with the binary outcome (dependent variable) of hazardous drinking measured just before graduation (T1) and again four years (T2) and 10 years (T3) after graduation. We performed logistic regression (LG) analyses to predict individual risk factors (measured at T1) for hazardous drinking four and 10 years later. The parameter risk estimates in LG are given as odds ratios (ORs) and represent probability; thus, they are generally easier to interpret compared with other parameter estimates (e.g., beta values from linear regression). We note that the weaknesses/limitations of LG include its sensitivity sample size, multicollinearity, outliers and loss of variance when dichotomizing variables with multiple response options. Multicollinearity is usually not a problem in larger samples such as ours. The highest correlation between our variables (in Paper I) was $r = 0.43$ (between neuroticism and reality weakness personality traits), which is acceptable for use in LG models.

We performed unadjusted (univariate) analyses of all predictor variables on hazardous drinking measured at T2 and T3, and the predictor variables were later entered into two adjusted (multivariable) models. To avoid type II errors, we included all predictors in the adjusted models that were significant at $P = 0.10$ in the univariate analysis. Only variables with $P < 0.05$ were considered significant in the adjusted models (211). In the first adjusted models of hazardous drinking at T2 and T3, all significant unadjusted predictor variables at T1 were included, except for baseline hazardous drinking. In the second model, we included all predictors from the first adjusted model, in addition to hazardous drinking measured at baseline (T1). This was done to control for baseline drinking and identify the adjusted variables that predicted a change in hazardous drinking during the observation period. Interaction analyses were also performed on gender, cohort and with all significant predictor variables.

**Paper II:** Paper II is a 15-year follow-up study for which we assessed predictors of hazardous drinking, beyond the individual-based risk factors identified in Paper I. We
investigated whether work stress, life stress and mental distress were independently associated with hazardous drinking after controlling for other individual factors. Since the dependent variable of hazardous drinking (binary) and other independent variables (e.g., work-related variables, life stress and mental distress) were measured repeatedly at different time periods (from T2 to T4) of these doctors’ careers, we used these longitudinal data. We performed generalized estimating equations (GEE) (212, 213) using an unstructured covariance matrix and robust estimator (the default giving more consistent estimates of covariance) to determine whether work stress, life stress or mental distress were associated with hazardous drinking. In addition, GEE were used to test the course of hazardous drinking and post hoc analyses. GEE are a method for dealing with repeated measures of categorical (binomial) outcomes that take into account dependency between the observation periods and include all available data for each individual participant. GEE take into account all observations, as opposed to older methods that exclude individuals with missing values. However, for GEE analyses to be valid, the strict assumption that missing values are missing completely at random must be met. GEE are an extension of the generalized linear model (mixed-method repeated-measures) used to analyse repeated measurements or other correlated or clustered data. Furthermore, GEE focus on mean effects and different working correlation matrices can be applied (e.g., independent, autoregressive, unstructured, etc.) As noted, the parameter risk estimates are reported as ORs, which are easy to interpret.

We performed unadjusted univariate analyses of the effects of all predictor variables measured at T1 and contextual variables and mental distress (measured at T2–T4) with hazardous drinking at T2–T4 as the dependent variable. Next, the variables were entered in three blocks to the adjusted analyses to isolate the effects of individual predictors from contextual work-related factors and mental distress (see Paper II, Table 2). Block 1 included the predictor (individual) variables measured at T1. Block 2 included the significant predictors from Block 1 and contextual variables (including life events) measured at T2–T4. The final model (Block 3) included significant predictors from Block 2 and mental distress measured at T2–T4 and time (T2) as reference. Variables with $P < 0.10$ were included in the blocks (adjusted model) to avoid type II error, but variables with $P < 0.05$ were considered significant in the final adjusted model (Block 3). We controlled for age, gender and cohort. Interaction analyses between gender and all other significant variables were also performed.

To determine any possible effects of doctors’ work, we repeated the analyses first with the contextual work-related factors in Block 1, then the life events in Block 2, then
mental distress in Block 3, and finally controlling for individual factors in Block 4. These analyses yielded the same results. Thus, we validated the findings of mental distress, life events and work-related stress as being independent of individual factors.

**Paper III:** To analyse life satisfaction (three-item scale), we used a linear mixed model (LMM) (214) with repeated unstructured covariance matrix and maximum-likelihood estimator (or algorithm) for all predictor variables. All independent variables, including the time, were treated as fixed effects. First, we performed unadjusted analyses of all predictor variables for life satisfaction from T2 to T4 as the dependent variable; these variables were then entered into two models in the adjusted analyses (see Paper III, Table 2). In Model 1, only significant predictors (i.e., individual) variables in the univariate analyses were entered. In the final Model 2, significant variables from Model 1, significant univariate contextual variables from Block 2, and time (T2 as a reference) were entered. We controlled for age, gender and cohort. Only variables with P < 0.05 were considered significant in the final adjusted Model 2. Additional interaction analyses using both gender and cohort were performed to identify any gender and/or cohort differences in the adjusted predictor effects.

We performed further analyses using the dependent variables from T2 to T4 with three clusters: (I) those with increasing life satisfaction (N = 71); (II) those with stable high life satisfaction (N = 331); and (III) those with decreasing life satisfaction (N = 87). LG analyses were performed on cluster III to determine the factors that predicted a decrease in life satisfaction (individual-related variables measured at T1, contextual work-related variables measured at T2).

### 3.9 Discussion of methodological issues

#### 3.9.1 Study design

This study used a longitudinal design, with measurements at four time periods during Norwegian doctors’ careers. The first observation was during their final year of medical school (average age 28 years, SD = 2.8) (T1), the second was four years after graduation (T2), the third was 10 years after graduation (T3), and the fourth and final observation was 15 years after graduation (T4). Most of the doctors in these samples completed their postgraduate or specialist training during this time period (177, 178) and may have established a family and developed an adult social role. This longitudinal design allowed us to follow the courses of outcomes (in our case, both hazardous drinking and life satisfaction).
over this 15-year period and to explore possible predictors and/or factors associated with these outcomes. However, we may have missed important information during this 15-year longitudinal study due to long gaps (i.e., of 4–5 years) between measurement time points. To achieve the most predictive values for life and work stress, no more than one or two years should elapse between measurement points (215). In addition, there may have been other variables beyond our focus that affected the relationships between predictors and outcome. These may include work-related variables such as practice venues. Therefore, we may have missed some associations and predictors due to the gaps between measurement points and unmeasured predictor variables. That said, to our knowledge, this is the most highly representative study of female doctors ever conducted, and it is therefore a study strength that we collected unique data with respect to women in medicine, from an international perspective previously unstudied. In Paper I, our independent variables were predictors insofar as they were measured at T1 (i.e., before the outcome variables measured at T2 and T3). In Paper II, the contextual factors (including life events and mental distress) were measured at about the same time as the outcome variable was measured, and are mean-level effects. Therefore, with respect to contextual factors, we cannot know which comes first. According to Hill’s criteria for causality assumption, temporality with respect to possible causation and identification of any risk factors was only fulfilled in Paper I (216). In Paper III, except for the personality variables, the other effects were mean level and measured at about the same time as the outcome variable was measured, However, in Paper III, we also performed LG analyses to see which factors predicted a decline in life satisfaction. In this case, the predictors of a decline in life satisfaction were measured before the outcome variable and are therefore ‘true’ predictors.

3.9.2 Selection bias

With regard to non-random attrition of those who responded at T1 but later dropped out of the study (n = 149), there were no significant differences in their the predictor variables at T1, with the exception of a slight age difference (roughly one year) between those who dropped out (mean = 28.5 years, SD = 3.3) and responders (mean = 27.7 years, SD = 2.8), indicating that slightly older participants were more likely to drop out of the study compared with younger participants (P = 0.006). There was no statistically significant difference between responders and non-responders with respect to association between age and drinking at T1 (r = –0.001 vs. r = –0.07, respectively). Nor were there statistically significant differences between responders and non-responders with regard to association between gender and
hazardous drinking at T1. No significant differences in point estimates of either independent or dependent variables were found at T1 between respondents who dropped out from T1 to T3 and those who responded at both time points. Slightly more males than females dropped out from T1 to T3 (31% vs. 25%) and drop-outs from T1 to T3 were significantly older than those who responded at both time points (28.4 vs. 27.5 years). Furthermore, more males than females dropped out from T1 to T4 (44% vs. 35%) and drop-outs from T1 to T4 were slightly but significantly older than those who responded at both time points (28.1 vs. 27.5 years). Male responder attrition means that we may have under-estimated the prevalence of hazardous drinking from T1 to T4, since this is consistently higher among men.

3.9.3 Reporting bias
This study included 30–40 pages of self-reported questionnaires, not clinical interviews. Doctors who did not respond to the questionnaires at all may have been more likely to have substance use disorders (including both alcohol and drug use) or have higher levels of stress and/or other mental disorders. Some authors have argued that self-report measures may lead to under-reporting of responders’ alcohol intake (217-219), while others contend that this type of bias is negligible in surveys of substance abuse so long as anonymity is assured (220). In sum, there is disagreement regarding the validity of self-reported questionnaires of alcohol intake. Nevertheless, our sample may have been at risk of under-estimating the prevalence of hazardous drinking.

Personality traits were not measured at baseline (T1) for the whole sample; rather these were measured in a random half of the Young Doctor Cohort at T1 and in the other half one year later (116). Despite a large longitudinal nationwide study showing changes over the entire lifespan (221), these traits show relatively high stability in individuals, with test–retest values of 0.5–0.8 over a four-year period (221, 222). Consistent with this, the test–retest value of general self-esteem (closely resembling neuroticism) in the Young Doctor Cohort measured at graduation and one year later was $r = 0.7$, indicating high stability (192).

3.9.4 Missing data
Respondents who completed all nine items on the AUDIT and all three items of life satisfaction were included in analysis. The same criteria applied to perceived job stress (emotional pressure, time pressure, fear of complaints and criticism, work–home interference), perceived job demands and autonomy, colleague support and perceived social support. Respondents who completed the questions about all 13 life event items and all five
items of the SCL-5 were also included in the analysis. For the personality dimensions, mean imputations were previously computed for individuals lacking ≤ 4 of the nine items on each dimension within the Young Doctor Cohort. This was done to reduce sample attrition due to the lack of responses on some personality items. It was not possible for us to recreate the personality variables without imputed subscales. It was therefore impossible for us to double-check if unimputed personality variables would give different results. However, the number of individuals with imputed scores was limited to 12 (reality weakness dimension), 19 (control dimension/conscientiousness), 20 (intensity dimension/extraversion), and 23 (vulnerability/neuroticism) out of 459 responders.

3.9.5 Psychometric considerations

Reliability (precision)

Reliability is the ability to obtain or reproduce the same results across different test administrations. In other words, reliability is how free the data obtained using the instrument (scale) are from random errors. Reliability also refers to the consistency and stability, or replicability, of the data. There are different methods for assessing the scale reliability: test–retest reliability, interrater reliability and internal consistency. We assessed the internal consistency of our scales/instruments using Cronbach’s alpha, which is based on the average correlation or intercorrelation between all items making up the scale. Cronbach’s alpha values > 0.70 indicate that more than 70% of the variance in an index is accounted for by the underlying (latent) phenomenon, and is generally considered acceptable (Table 4).

With regard to our independent variables, Cronbach’s alpha values for the personality dimensions of control or conscientiousness (Cronbach’s alpha = 0.63) and reality weakness (Cronbach’s alpha = 0.62) were < 0.70 and thus considered quite low (i.e., questionable). The other personality dimensions, vulnerability (neuroticism) and intensity (extraversion/introversion), had acceptable Cronbach’s alpha values > 0.70, reducing our risk of type II errors (i.e., false negatives) (50). Low Cronbach’s alpha values are caused by lower correlations between an instrument’s items due to random errors. The variables of perceived social support and mental distress, and the dimensions of perceived job stress (including work–home interference), perceived job demands and colleague support had Cronbach’s alpha values ranging from 0.71 to 0.91, which are considered acceptable to excellent, and thereby reducing our risk of type II errors. The Cronbach’s alpha for our dependent variable of life satisfaction was very good (0.82–0.83), indicating high internal consistency and low
risk of type II errors. A major strength of this longitudinal study is that our work-related factors were measured repeatedly, which also increases the reliability of our findings. Many factors can affect scale reliability, including the degree to which the instrument has been tested, the evaluators using the instrument (i.e., their experience in using the instrument), data analysis approaches and variation in the material (223). Internal consistency may also be boosted by systemic errors such as set response and scale effects (i.e., consistently using the same response on a given type of response scale), which can artificially increase correlations between measures using the same type of items or response scales, or even increase the risk of type I errors (false positives).

**Table 4. Cronbach’s alpha values**

<table>
<thead>
<tr>
<th>Cronbach’s alpha</th>
<th>Internal consistency reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥ 0.90</td>
<td>Excellent</td>
</tr>
<tr>
<td>0.80–0.89</td>
<td>Good</td>
</tr>
<tr>
<td>0.70–0.79</td>
<td>Acceptable</td>
</tr>
<tr>
<td>0.60–0.69</td>
<td>Questionable</td>
</tr>
<tr>
<td>0.50–0.59</td>
<td>Poor</td>
</tr>
<tr>
<td>&lt; 0.50</td>
<td>Unacceptable</td>
</tr>
</tbody>
</table>

**Validity (accuracy)**

Validity is the accuracy of an instrument, or the extent to which the instrument measures what it is intended to measure. Reliability is a necessary but insufficient precondition for validity. Data reliability does not necessarily mean that they are also valid. There are several types of validity: content validity, criterion validity (concurrent and predictive), and construct validity.

*Content validity* is whether the items on the instrument adequately represent the construct we want to measure. *Criterion validity* refers to how well, or poorly, an instrument correlates mathematically with another well-accepted instruments (a so-called ‘gold standard’) purported to measure the same phenomenon at present (*concurrent validity*), in the future (*predictive validity*) or in the past (*retrospective validity*). When there is no well-accepted instrument or gold standard, *construct validity* can be assessed by observing whether the instrument performs well according to our hypotheses. The methods for this approach include finding strong correlations with measures of similar constructs (*convergent validity*) and low/no correlations with theoretically unrelated measures (*divergent validity*).
Most of the instruments used in this thesis were validated in previous research, although we performed additional analyses to further validate or substantiate our dependent variables of hazardous drinking, both the single-item measure and our cut-off score on the AUDIT-9. Our dependent single-item variable for hazardous drinking was based on a cut-off at ‘2–3 times or more often per month’, which has previously been used with students and young adults, but less often with a middle-aged population. Thus, we assessed its construct validity using self-prescribed medications, as described under Materials and methods. Based on these findings, we regard our validity to be sufficient. To avoid under-estimation of hazardous drinking among women, we validated our single-item measure of hazardous drinking among women with the item ‘having four drinks administered on the same occasion’, and found a strong correlation ($r = 0.83$, $P < 0.001$), indicating high validity.

The validity of our cut-off score for hazardous drinking measured with the AUDIT-9 was also assessed against self-prescribed medications and also successfully indicated validity. Our three-item measure of life satisfaction has not been previously validated except for the first, single-item version, which has been used and validated previously (20, 135). However, as noted, different terminologies and aspects of physician wellness have been used in the literature; we believe that our three-item measure of life satisfaction includes both the evaluative and hedonic aspects of well-being. The reliability of this measure was consistently good at several time points, with Cronbach’s alpha values of 0.82–0.83 (T2–T4). Using a three-item measure instead of a single-item measure increased the reliability of our findings, although future studies will be needed to further validate this measure of life satisfaction.

3.9.6 External validity
The representativeness of the NORDOC data is described in detail in the study and sample design section under Materials and methods. The response rate in our study was much higher compared with most studies of doctors, indicating that it is highly representative. About 90% of our target sample responded at least once, despite our questionnaire battery being 30–40 pages long at each follow-up over a 15-year span. Low survey response rates are not uncommon among samples of doctors, with a recent survey of US doctors reporting that only 26.7% responded (25). Although doctors’ work environments, hours and conditions may differ across European countries and from the USA, they are similar across Scandinavian countries. The highly regulated working hours and nearly identical social benefits across these countries mean that our findings can likely be generalized to other Scandinavian countries, and possibly to other countries in Northern Europe. The NMA strongly asserts that
Norwegian doctors should work no more than an average of 40–45 hours/week. In other parts of Europe, and around the world, much higher working hours can be found. Doctors’ working conditions also differ across continents; thus, in the present study we included individual vulnerability factors to determine those that have independent (or adjusted) effects on our dependent variables. We anticipate that these findings may be more readily generalizable to other countries, since they are less dependent on working conditions and environment. With respect to some work-related factors in Paper III, we expect that they would have an even stronger effect in the USA, where working hours and conditions are less regulated.

Furthermore, the effects of work–home stress may be more pronounced in a country with worse social security benefits and less coverage of childcare than in the Scandinavian countries.

3.9.7 Ethics

This study was conducted according to the guidelines of the Regional Committee for Medical Research (approval references REK 2010/788a and 2013/1585) and was approved by the Norwegian Data Inspectorate. All participants provided informed consent, and respondents’ identities were not divulged to the researchers by the Central Bureau of Statistics Norway (i.e., the researchers only had access to anonymous data).
4. SUMMARY OF PAPERS

4.1 Paper I

Risk factors measured during medical school for later hazardous drinking: a 10-year, longitudinal, nationwide study (NORDOC)

Javed Iqbal Mahmood, Kjersti Støen Grotmol, Martin Tesli, Per Vaglum, Reidar Tyssen

Aims: This study investigated the prevalence of hazardous drinking, its course over time and the risk factors measured in medical school for future hazardous drinking among doctors.

Methods: Two cohorts of graduating medical students (N = 1052) from all four Norwegian universities (NORDOC) were surveyed at graduation (T1) and again four (T2) and 10 years later (T3). Longitudinally, 53% (562/1052) of the sample responded at all three time points. Hazardous drinking was measured as ≥ 5 drinks on one occasion, at least 2–3 times/month. Predictor variables of hazardous drinking assessed in LG models, after controlling for cohorts, included parental history of alcohol problems, having children, religious activity, use of alcohol to cope with tension and personality traits.

Results: The prevalence of hazardous drinking declined significantly from T1 (14%) to T2 (10%) (P = 0.001), but not from T2 to T3 (8%). Significant adjusted predictors for hazardous drinking at T2 were male gender (OR = 2.0, 95% CI = 1.01–4.1, P = 0.04), use of alcohol to cope with tension (OR = 2.2, 95% CI = 1.05–4.8, P = 0.03) and hazardous drinking at T1 (OR = 9.8, 95% CI = 4.9–19.5, P < 0.001). Significant adjusted predictors for hazardous drinking at T3 were older age (OR = 1.1, 95% CI = 1.02–1.2, P = 0.01), male gender (OR = 3.6, 95% CI = 1.6–8.2, P = 0.002) and hazardous drinking at T1 (OR = 7.5, 95% CI = 3.4–16.7, P < 0.001). No significant gender or cohort interactions were found with any of the significant predictors; thus, the main effects were similar between men and women and between the two cohorts. About 23% (N = 18) of those drinking hazardously at T1 continued to drink hazardously at both T2 and T3.

Conclusion: Hazardous drinking and drinking to cope with tension in medical school were both significant predictors of later hazardous drinking and should therefore be targeted in preventive interventions during medical school.
4.2 Paper II

Contextual factors and mental distress as possible predictors of hazardous drinking in Norwegian medical doctors: a 15-year longitudinal, nationwide study

Javed Iqbal Mahmood, Kjersti Støen Grotmol, Martin Tesli, Per Vaglum, Reidar Tyssen

Aims: To examine whether contextual factors (work and life stress) and mental distress are independently linked to hazardous drinking over the first 15 years of doctors’ medical careers, after controlling for individual predictors.

Methods: Two nationwide cohorts of Norwegian doctors (N = 1052) graduating from all four Norwegian universities were surveyed during their final year of medical school (1993/94 and 1999, respectively) (T1), and then four (T2), 10 (T3), and 15 (T4) years later. Hazardous drinking was measured using the nine-item version of the AUDIT, and the results were analysed using GEEs. Work-related variables, life events and mental distress were measured at T2–T4. Individual predictors included personality traits and having no religious activity.

Results: Ninety per cent of the sample (947/1052) responded at least once and 42% (450/1052) responded at all four time points. Hazardous drinking was reported by 16% at T1, 14% at T2 and T3 and 15% at T4. There was a significant decline in hazardous drinking from T1 to T2 (OR = 0.78, 95% CI = 0.62–0.97, P = 0.029). Life events (OR = 1.16, 95% CI = 1.03–1.30, P = 0.009) and mental distress (OR = 1.43, 95% CI = 1.14–1.80, P = 0.002) were adjusted predictors of hazardous drinking, in addition to male gender (OR = 1.67, 95% CI = 1.14–2.43, P = 0.008), having no religious activity (OR = 1.63, 95% CI = 1.02–2.61, P = 0.040), drinking to cope with tension (OR = 2.77, 95% CI = 1.74–4.39, P < 0.001) and conscientiousness personality trait (OR = 0.84, 95% CI = 0.77–0.93, P = 0.001). A significant interaction was found between gender and mental distress (OR = 1.48, 95% CI = 1.03–2.14, P = 0.034) and between gender and life events (OR = 1.20, 95% CI = 1.06–1.36, P = 0.004) indicating that mental distress and life events were more important predictors of hazardous drinking among men than they were among women.

Conclusions: Mental distress, life events and drinking to cope were more important risk factors for hazardous drinking than was work-related stress. Preventive measures should be taken to reduce mental distress (i.e., anxiety and depressive symptoms) and drinking to cope with tension among doctors. The absence of an effect of work-related factors might be attributed to a highly regulated Scandinavian work life.
4.3 Paper III

Life satisfaction in Norwegian medical doctors: a 15-year longitudinal study of work-related predictors

Javed Iqbal Mahmood, Kjersti Støen Grotmol, Martin Tesli, Torbjørn Moum, Ole Andreassen, Reidar Tyssen

Aims: The aims of this study were to determine the long-term, work-related predictors of life satisfaction among established Norwegian doctors and to identify these predictors in a subgroup of doctors who reported a decline in their life satisfaction.

Methods: Two nationwide cohorts of doctors (N = 1052) were surveyed at graduation from medical school (1993/94 and 1999, respectively) (T1), and then four (T2), 10 (T3) and 15 (T4) years later. Work-related predictors of life satisfaction (three items) obtained at T2–T4 were analysed using mixed-model repeated-measures analyses. We controlled for individual (e.g., personality) and lifestyle (i.e., physical training and drinking behaviours) confounders. LG analyses were used to identify predictors of a decrease in life satisfaction.

Results: Ninety per cent (947/1052) responded at least once, and 42% (450/1052) responded at all four times. Work-related predictors of higher life satisfaction in the adjusted model were work–home stress ($\beta = -0.20, 95\% \text{ CI} = -0.25–0.16, P < 0.001$), perceived job demands ($\beta = -0.10, 95\% \text{ CI} = -0.15 to -0.05, P < 0.001$) and colleague support ($\beta = 0.05, 95\% \text{ CI} = 0.04–0.07, P < 0.001$). The new adjusted individual predictors identified included female gender ($\beta = 0.09, 95\% \text{ CI} = 0.003–0.18, P < 0.043$), reality weakness trait ($\beta = -0.03, 95\% \text{ CI} = -0.06 to -0.003, P = 0.029$), use of alcohol to cope with tension ($\beta = -0.24, 95\% \text{ CI} = -0.37 to -0.12, P < 0.001$) and hazardous drinking ($\beta = -0.16, 95\% \text{ CI} = -0.27 to -0.05, P = 0.003$). A significant positive interaction was found between gender and drinking to cope ($\beta = 0.28, 95\% \text{ CI} = 0.06–0.50, P = 0.011$), indicating that the negative effect of drinking to cope was significantly stronger among males. In addition, the negative effects of neuroticism ($\beta = 0.04, 95\% \text{ CI} = 0.004–0.09, P = 0.032$) and perceived job demands ($\beta = 0.09, 95\% \text{ CI} = 0.02–0.16, P = 0.005$) on life satisfaction were stronger among men than they were among women.

There was also a significant cohort effect of work–home stress ($\beta = -0.07, 95\% \text{ CI} = -0.14 to -0.01, P = 0.016$) on life satisfaction, indicating a significant stronger effect of this stress among the younger cohort of doctors, who graduated in 1999. The only adjusted predictors of a decrease in life satisfaction over time were the personality trait neuroticism (OR = 1.25,
95% CI = 1.07–1.47, P = 0.004) and colleague support (OR = 0.84, 95% CI = 0.73–0.97, P = 0.021).

**Conclusions:** Work–home stress, perceived job demands and colleague support were the most important work-related predictors of life satisfaction among this sample of doctors. When controlling for personality traits, female doctors were more satisfied with their lives than were male doctors. In addition to neuroticism, colleague support was an independent predictor of a decline in life satisfaction during these years. These findings suggest that improving work-related factors by targeted interventions, including providing a supportive work environment, may increase life satisfaction among doctors.

### 4.4 Additional findings (not previously published)

**Characteristics of individuals who reported drinking hazardously at T1 and who continued this behaviour at T2 and T3**

Twenty-three per cent (N = 18) of all hazardous drinkers (N = 79) at medical school (T1) were drinking hazardously four (T2) and 10 years later (T3). We performed additional analyses to examine the characteristics of continuous drinkers. Almost all of these were men (N = 16/18). LG analyses showed that the adjusted predictors of continuous drinkers were male gender (OR = 8.33, 95% CI = 1.86–37.26, P = 0.006), use of alcohol to cope with tension (OR = 2.91, 95% CI = 0.97–8.73, P = 0.056) and conscientiousness (OR = 0.73, CI = 0.53–1.003, P = 0.052). The only significant predictor was male gender, but there was a trend towards significance for both drinking to cope with tension and conscientiousness personality trait. We expect that the lack of significant effects here are from type II errors because these samples were relatively small (N = 18 and 79).

**Comparison of doctors in surgical specialties versus non-surgical specialties with respect to hazardous drinking (nine-item version of the AUDIT)**

Additional post hoc analyses at T4 with the NORDOC sample were performed among doctors working in surgical specialties (including anaesthesiology) versus non-surgical specialties, with respect to hazardous drinking. Working in surgical specialties was positively associated with hazardous drinking (OR = 1.91, 95% CI = 1.07–3.40, P = 0.028). When separate analyses were performed for men and women, we found no significant difference in hazardous drinking among females (N = 38) in surgical specialties versus non-surgical specialties (OR = 1.37, 95% CI = 0.52–3.59, P = 0.511), whereas in men (N = 51) there was
clear trend towards more hazardous drinking in surgical specialties (OR = 2.12, 95% CI = 0.99–4.52, P = 0.051), although this was not significant.
5. GENERAL DISCUSSION OF THE RESULTS

5.1 Prevalence of hazardous drinking among doctors as a group

The prevalence of hazardous drinking among doctors, measured with a single-item question, was described in the results above. Using the AUDIT-9, the prevalence rates of hazardous drinking were 16% (24% for men, 11% for women, \( P = 0.002 \)) at graduation; 14% (21% for men, 9% for women, \( P = 0.004 \)) four years after graduation; 14% (21% for men, 10% for women, \( P = 0.003 \)) 10 years after graduation; and 15% (22% for men, 10% for women, \( P = 0.002 \)) 15 years after graduation. In addition, men reported more hazardous drinking than did women at each observational time period, regardless of the item/instrument used to measure hazardous drinking.

A direct comparison of hazardous drinking prevalence estimates with other studies is difficult due to their use of different instruments and/or cut-off levels. In addition, different alcoholic beverages contain different quantities of pure alcohol, which also differ across countries. For instance, one standard drink contains 12 g of pure alcohol in Norway, whereas it contains 14 g in the USA, 10 g in Australia, and 20 g in Austria and Japan. In the UK, one standard drink (UK unit of alcohol) contains 8 g of pure alcohol (224). These discrepancies make it difficult to compare the exact prevalence estimates of hazardous drinking across studies. A prospective UK study showed that 20% (25% for men, 16.9% for women) of final-year medical students and 18% (12.9% for men, 20.8% for women) of those one year post-graduation were involved in binge drinking (i.e., drinking over half the recommended ‘safe limit’ per week during one session, which is ≥ 10 units for men and ≥ 7 units for women) (39). Despite the UK being a distinct drinking culture compared with Norway, that study also showed a reduction in the prevalence of binge drinking after leaving medical school. This may be due to increased occupational duties and family responsibilities. In another US cohort recruited from 16 medical schools, 37% (43% of men and 24% of women) of final-year medical students drank excessively (i.e., consumed ≥ 5 drinks on at least one occasion within the previous month or > 2 drinks/day on average for men or > 1 drink/day on average for women) (76). Thus, it seems that our prevalence estimates of hazardous drinking among final-year medical students in Norway are lower than those reported among medical students in the UK and USA. This might be explained by the different drinking cultures across countries, the high prices of alcoholic beverages in Norway, in addition to restrictive control policies in Norway on alcohol sales.
With regard to hazardous drinking among doctors after medical school, a recent nationwide Danish survey showed a rate of 18.3% (58.4% men), which is much higher than in our sample, although that study’s response rate was 48.6% (62). The Danish estimate may not be surprising given the differences in drinking cultures and that alcohol regulation in Denmark is more liberal compared with Norway and Sweden (225).

Regarding the general Norwegian population, a recent report showed that about 15% (20.9% of men, 9.0% of women) of the adult population sampled during 2015–2017 engaged in risky alcohol consumption based on self-reported AUDIT scores (using a cut-off ≥ 8 for both genders) (226). The highest level of risky drinking was reported among men and risk reduced with age. This prevalence among the Norwegian adult population is consistent with the present study, meaning that Norwegian doctors drink at about the same level as does the general population. Like Sweden, Norway is temperate in terms of alcohol consumption, and the Norwegian government enforces strict public policies limiting alcohol consumption based on age (i.e., 18 years for beer and wine, 20 years for brandy/spirits), with high costs and limited access (e.g., reduced opening hours, fewer selling points, marketing forbidden). These factors may also affect doctors’ drinking. However, most practising doctors in Norway have an income above the population mean and may therefore be less affected by the high prices of alcohol. Among the general Norwegian population during 1993–2000, higher education and higher income levels predicted alcohol consumption (227). Interestingly, we have no indication that levels of drinking are lower among doctors than among other Norwegians in the same socio-economic group. This is consistent with a study of Norwegian doctors conducted 30 years ago, which showed the same level of hazardous drinking among doctors of both genders and others above 45 years of age, with even higher levels of drinking frequency among female doctors compared with other Norwegians (32). Unfortunately, we lack recent population studies with which we can compare Norwegian doctors.

5.2 Stability of hazardous drinking among doctors as a group
There was a significant decline in the prevalence of hazardous drinking from medical school to four years post-graduation, measured by both a single item and the nine-item AUDIT; thereafter, the prevalence remained stable throughout the 15-year follow-up. This initial decline in the prevalence of hazardous drinking is in keeping with other follow-up and cross-sectional studies of the general population (79, 228) and in Norway (229), which generally show a decline in hazardous drinking in individuals during and after their 20s. A recent
epidemiological study in the UK also examined life course trajectories of alcohol consumption using longitudinal data from nine cohort studies and found that mean alcohol consumption peaked at around age 25 years, and then declined and plateaued during mid-life, before declining again around age 60 years (230). This decline in alcohol consumption after their 20s might be explained by individuals having established an adult social role, with a partner and/or profession, as well as parenting (229, 231, 232).

5.3 Individual predictors of hazardous drinking

In the present study, male gender, use of alcohol to cope with tension, older age, low level of conscientiousness, having no religious activity and hazardous drinking in medical school were all independent risk factors for later hazardous drinking. This is consistent with previous studies of both doctors and the general population. We now discuss these findings in relation to the present study and our additional analyses.

That male gender significantly predicted hazardous drinking is not a novel finding. As noted, a Norwegian study reported higher rates of hazardous drinking among female surgeons than among female non-surgeons (18% vs. 7.6%, respectively); however, when separate gender analyses were performed, being a surgeon was a significant predictor of hazardous drinking for both females and males (60). In our additional analyses, we compared doctors in surgical specialties vs. non-surgical specialties with respect to hazardous drinking, and found that working in surgical specialties was positively associated with hazardous drinking. When gender analyses were performed, we found the expected trends but no significant differences. This was presumably due to the relatively small samples that may have led to type II errors. Based on these cumulative findings, we might conclude that working in surgical specialties is a risk factor for hazardous drinking in the NORDOC sample. The increasing proportion of female doctors (233, 234) and other studies showing a narrowing of the gender gap with respect to drinking (235, 236) may explain the higher rates of hazardous drinking among female surgeons vs. female non-surgeons, as reported by Rosta and Aasland (60).

As we hypothesized, the use of alcohol to cope with tension, measured during medical school (T1), significantly predicted hazardous drinking throughout the 15-year follow-up, even after controlling for other factors that predict alcohol problems. This is an important finding. In a previous NORDOC study, our research group reported that this variable predicted hazardous drinking six years post-graduation among male doctors (10).
However, the present study is the first to demonstrate that drinking to cope predicts hazardous drinking throughout the 15 years post-graduation. This finding is consistent with another longitudinal study of the general population that concluded that baseline drinking to cope was associated with more alcohol consumption and drinking problems across all four observations over a 10-year period (237). In the present study, drinking to cope in medical school translated to a nearly threefold increased risk (OR = 2.77) of hazardous drinking over the 15-year follow-up. Doctors may have learned this unhealthy drinking behaviour during medical school and some may continue this prominent behaviour in stressful situations after graduating.

When hazardous drinking was measured with a single item, drinking to cope in medical school was a significant predictor of hazardous drinking four years after graduation, but not 10 years after graduation. This may be because the single-item measure of hazardous drinking is less reliable, or it may be that doctors are more prone to drinking to cope during their earlier career (i.e., during residency, internship, or postgraduate training) because they have more work-related stress during these phases of their careers (42, 238). Alternatively, doctors may drink to cope with mental health problems, particularly anxiety and depressive symptoms, which have been reported to be highest during their early career phases (43, 140).

We did not find a statistical association between drinking to cope with tension and gender, meaning that the predictor effect of drinking to cope was similar for men and women. The prevalence of drinking to cope was also equal between males and females in a previous Norwegian cross-sectional study of police and ambulance personnel (103), and in our group’s previous NORDOC study of medical students (9). Because most alcohol risks are gender-, and specifically male-dependent, drinking to cope may therefore represent a particular risk for drinking problems among women, at least among those in the same socio-economic group as physicians.

**Older age** in medical school was a significant predictor of hazardous drinking 10 years post-graduation. For each year of age, there was a 15% increased risk of hazardous drinking during students’ final year of medical school. The admission scheme of Norwegian medical schools gives students aged over 20 years two age points/year (up to 8 points) without requiring relevant work or education. A previous NORDOC study showed that older medical students were less likely to drink to intoxication but were more likely to drink to cope with
tension (9). Those who are older when in medical school may belong to a unique group that should be further studied.

The personality trait of low conscientiousness was an independent predictor of hazardous drinking throughout the 15 years of follow-up, even after controlling for mental distress and other factors. The long-term predictive validity of low conscientiousness (or greater impulsiveness) is a novel finding among Norwegian doctors, although we previously found this trait to be an independent predictor of hazardous drinking among medical students during a six-year follow-up study (10) and is in keeping with other population cohort studies (90).

We did not find any effect of personality traits on hazardous drinking on LG analyses using the single-item measure of hazardous drinking; however, with GEE using the AUDIT-9, the personality trait of low conscientiousness was a significant predictor. This may be due to our use of more powerful statistics and a more reliable measure of hazardous drinking. A cross-sectional study of Australian doctors showed that the personality traits of neuroticism and extraversion were associated with hazardous drinking, although this retrospective study was also hampered by potential selection bias from a relatively low response rate (36%) (37). A large representative cross-sectional NORDOC study of medical students showed that high general self-esteem (which resembles low neuroticism) is linked to hazardous drinking in adjusted models (9). Extraverts are usually more sociable and impulsive, and use risky behaviours to seek stimulation and arousal. Extraversion has also been linked to problematic drinking in the general population (90, 239), although in the present study extraversion was a significant predictor of hazardous drinking only in unadjusted analyses. This may be because extraversion and conscientiousness were negatively correlated ($r = -0.15, P < 0.001$); additional analyses showed that conscientiousness suppressed the effect of extraversion. In summary, our cohort and those in other reports show that conscientiousness personality trait is a relatively strong predictor of hazardous drinking.

**Having no religious activity** during medical school was an independent predictor of hazardous drinking throughout the 15-year follow-up when hazardous drinking was measured with the most reliable measure (AUDIT-9). Religiosity as a protective factor against hazardous drinking might be explained by social and cultural norms and individuals’ attitudes towards drinking (240-242). This is not an original finding, although the relationship between religiosity and drinking has been studied less thoroughly among doctors. To our knowledge,
only one prospective study from 1990 has shown a lack of religiosity to predict mid-life alcohol abuse (38). However, this finding is in line with studies of college students (81, 82, 243) and of the general population (83, 244). Overall, religiosity seems to be a protective factor against hazardous drinking, which is interesting in a relatively secular, modern Scandinavian country.

**Hazardous drinking** in medical school was an important risk factor for hazardous drinking even 10 years post-graduation. This is not a novel finding, although to our knowledge, no recent long-term or prospective studies have assessed the risk factors of drinking in medical doctors. A 2008 prospective study of the general population showed that risky drinking 1–3 times/month doubled the risk of developing alcohol use disorder, whereas this risk quadrupled among those engaged in risky drinking ≥ 3 times/week (8). Thus, hazardous drinking should not be considered a transient or ‘state’ phenomenon, but may instead be considered a ‘trait’ behaviour. Those who exhibit this drinking pattern in medical school may be at greater risk of developing serious alcohol problems during later life, which raises the possibility of concomitant problems practising medicine.

### 5.4 Additional findings as predictors of enduring hazardous drinking

Additional analyses showed that individuals who engaged in hazardous drinking at all three observational time periods were almost exclusively men. Unfortunately, we could not identify significant predictors in this group of hazardous drinkers, likely due to the very small sample (i.e., leading to type II errors). Nevertheless, the trends in the effects of drinking to cope and low conscientiousness were in keeping with the significant findings from our larger sample, highlighting the importance of focusing on this behaviour and personality trait as potential risk factors present at the beginning of a doctor’s medical career. We believe this subgroup may be at risk of developing more serious alcohol disorders and is an important target group for preventive measures.

### 5.5 Independent factors related to hazardous drinking

We did not find any effect of work-related factors or other psychosocial factors on hazardous drinking, but mental distress and life stress were independently associated with hazardous drinking in this 15-year follow-up study of Norwegian doctors. Cross-sectional studies among US doctors have shown alcohol abuse or dependence to be associated with burn-out, depressive symptoms and suicidal ideation (25, 26), although the direction of these
relationships is unclear from such studies. In addition, these two studies reported that the longer the doctors practised, the less likely they were to have alcohol abuse or dependence, possibly representing a so-called ‘healthy worker effect’, and a similar association was found with hours worked and nights on call, with longer hours worked and more nights on call lowering the prevalence of alcohol abuse or dependence (25, 26). Previous research is inconclusive with regard to work-related factors and drinking. Doctors in Germany (150) and the UK (245) have also reported working longer hours than their Norwegian colleagues, although to our knowledge, no German or UK longitudinal studies have examined the associations between work-related stress and drinking among doctors. However, we located one cross-sectional study of German veterinarians which examined the association between psychosocial stress and alcohol consumption (246). Psychosocial stress was found to be a risk factor for binge drinking (≥ 5 drinks on at least one single occasion during the previous month) and regular drug use. The main factors underlying psychosocial stress were time pressure due to heavy workload, difficulties balancing one’s professional and personal lives, dealing with difficult customers, insufficient free time and long working hours. Other cross-sectional studies have found no relationship between occupational stress and drinking among either Norwegian police or Australian dentists (103, 104), which is consistent with the present study.

In conclusion, and in line with other research, we failed to find a direct link between working hours, or any other work-related factors, and hazardous drinking. Prospective longitudinal studies of doctors in other countries are needed, in locations with fewer working hour regulations and different work-life contexts. It would also be of interest to determine whether drinking to cope mediates work stress and hazardous drinking. We did not control for any mediators in the present study (e.g., drinking to cope with tension) because our work-related factors were not statistically significant, even at a univariate level.

Mental distress (i.e., anxiety and depressive symptoms) had an independent effect on hazardous drinking after controlling for individual and contextual work-related factors, indicating a co-occurrence of, and possible risk for, comorbid mental and alcohol disorders. This is not an original finding, but one that should be emphasized given its greater importance than contextual work-related stress and its relative commonality among practising doctors. That doctors are vulnerable to anxiety and depression also supports the notion that they are not invincible (247), and doctors’ health care should therefore also be prioritized.
Low threshold initiatives such as the Villa Sana scheme and other tailored health services in Norway are hugely important (248, 249).

The prevalence estimates of comorbid mental and substance use disorders (dual diagnosis) among doctors have been barely studied in representative samples. However, a Spanish study of doctors admitted to a psychiatric in-patient unit for health professionals reported that 71% had at least one diagnosed substance use disorder. Among those reporting a substance use disorder, 29% had comorbid conditions (250). In that study, 60% of the doctors with dual diagnosis had affective disorders (47% had major depression, 13% had bipolar disorder), 10% had anxiety and 5% had psychotic disorders. The most prevalent substance used disorders were alcohol dependence (43%), alcohol dependence with another substance use disorder (28%) and benzodiazepine dependence (17%). The most common comorbidity among doctors with a dual diagnosis was major depression and alcohol dependence (25%). It is important that doctors seek help early to combat mental illness and drinking problems, before these become advanced; studies have shown good prognosis for doctors in treatment.

In the present study, we did not aim to assess whether mental distress causes hazardous drinking, or whether hazardous drinking is caused by mental distress (in other words, the causal direction). Rather, the aim was to determine whether there is an association between work-related stress and/or mental distress and hazardous drinking over time, after controlling for other possible individual-based risk factors. The literature is inconsistent regarding the relationship direction between alcohol use disorders and depression (251). Therefore, further research is needed to examine this relationship, and whether it is indeed causal or attributable to common, shared risk factors such as genetic susceptibility (252).

**Life events** had an independent effect on hazardous drinking among Norwegian doctors, which is consistent with findings among the general population (98). However, our post hoc univariate analyses showed that life events associated with hazardous drinking among men were as follows: partner being unemployed/granted leave; doctor himself/herself or a close family member having been involved in a serious violation of the law; and moving away from one’s parents.
5.6 Level of life satisfaction among Norwegian doctors (longitudinal sample)

Levels of life satisfaction among Norwegian doctors did not differ significantly across the three observational time periods: 4.55 (SD = 0.82) at T2; 4.60 (SD = 0.81) at T3; and 4.62 (SD = 0.81) at T4. In addition, there were no significant gender differences at any of these time periods. This indicates a relatively stable trend on the group level from years 4 to 15 after leaving medical school (T2–T4), in both genders. Thus, it appears that changes in career and life circumstances during Norwegian doctors’ initial training and postgraduate years have little impact on their overall well-being. A previous NORDOC cohort study showed that total work stress changed little during the first 10 years of doctors’ careers, although three work stress factors declined and work–home interference increased, the latter being the most important predictor of life satisfaction (114). The findings from our study are consistent with a longitudinal study during 1994–2002 of 1,174 Norwegian doctors, which showed high general life satisfaction (142). Other factors, such as personality and social support may dampen the impact of stress during doctors’ initial years of practising.

In contrast to our study, a cross-sectional US study reported that older doctors (average age 43 ± 10 years) in both academic and clinical careers were more satisfied with their work and life compared with younger doctors (161). This may be due to the heavier workloads among younger doctors in postgraduate training (i.e., residency) in the USA compared with Scandinavia (120). A series of prospective longitudinal Swiss studies of doctors in residency showed lower levels of life satisfaction for both genders compared with the general population (143, 144), although they also reported domain-specific life satisfaction and are therefore difficult to compare with ours.

In conclusion, and in agreement with a previous longitudinal study of Norwegian doctors, life satisfaction levels did not change significantly from years 4 to 15 post-graduation. Epidemiological studies will be needed to determine whether Norwegian doctors’ life satisfaction differs from that of the general Norwegian population. Interestingly, a previous NORDOC study showed lower life satisfaction among Norwegian doctors during their postgraduate years compared with that among a general population sample (20); it is unknown whether this difference increases across doctors’ mid-career years.

5.7 Independent work-related factors associated with life satisfaction

With respect to work-related factors, we found that work–home stress, perceived job
demands and colleague support all had independent impacts on doctors’ life satisfaction. Thus, these factors appear to be important to doctors working in Scandinavia.

In the present study, doctors with low work–home stress reported higher well-being, which is consistent with cross-sectional studies of doctors (148, 253). Work–home stress also seems to be highly prevalent among hospital doctors in Germany, who reported higher work–home stress compared with the general German population (253). A previous NORDOC study found that reducing work–home stress was an independent predictor of improved job satisfaction among Norwegian doctors (254). In another NORDOC study, work–home stress predicted doctor burn-out (210). US studies have also concluded that work–home stress explains much of the variance in burn-out, as well as explaining depressive symptoms (255, 256). All these studies corroborate our findings about the importance of work–home stress among doctors.

Some studies have reported work–home stress to be more impactful for female doctors than for men (253, 255, 257); however, we did not find a gender difference between the doctors in our sample. This may be explained by divergent social roles with respect to work–home stress, since females traditionally have the primary responsibility for childcare and domestic duties. Moreover, the lack of a significant gender difference in our study with respect to work–home stress could also be due to greater social support for women, gender equality, and men and women sharing parental roles in Scandinavia. In Norway, fathers are now obliged to take 15 weeks of the 12 months of paid parental leave before their child reaches three years of age, otherwise the parents lose this benefit (258). A recent qualitative study showed that compared with any gender difference, there is a greater difference between junior and senior doctors with respect to working long hours and work–life balance, with older doctors tending to view their work as a lifestyle and younger doctors preferring fixed working hours (259). We suspect that some work-related factors identified in Scandinavia would be even more significant in countries with predominantly private health care systems, such as in the USA, where doctors have higher workloads and less regulated working hours. Thus, these findings may also have greater validity outside Scandinavia.

A low level of perceived job demands was independently related to higher well-being among Norwegian doctors. This is consistent with a previous cross-sectional study showing that doctors who have greater control over their work, have reasonable job demands and have greater colleague support also report higher satisfaction, organizational commitment and
psychological well-being compared with doctors with less perceived control, greater stress from uncertainty, higher job demands and less social support, who were at greater risk for burn-out (260). Despite the various studies that use the demand–control model in occupational health (261-263), few have focused on its association with general well-being and life satisfaction among doctors.

A cross-sectional Swedish study found that male and female general practitioners with high work demands and low control had lower well-being and impaired health compared with a reference group (264). That study also showed a stronger association between low control over work and lower well-being among male general practitioners compared with that among female general practitioners. In 2016, Eckleberry-Hunt and colleagues found that it was not the number of working hours/week, but rather the perception of workload that was associated with lower physician wellness (17). This is consistent with previous NORDOC studies that have all failed to find a direct association between working hours and mental health (116, 207). We believe that the effects on life satisfaction among doctors may derive from some specific job demands uniquely related to doctors’ work life. Such demands may include, for example, having responsibility for the medical care and treatment of severely ill patients. In addition, hospitals now seem to put more pressure on individual doctors by expecting increased productivity. Doctors express increased frustration over the greater number of administrative tasks and desk-work occupying their working hours, often at the expense of clinical work and patient care (265, 266). Quality of care assessments are often made by the doctors themselves, despite the fact that much of this work could be delegated to administrative staff. For example, in the USA, inefficient electronic health records steal time and create a great deal of frustration (267). In sum, the independent effects of job demands in our study may capture some of issues particular to doctors’ work situations. Interestingly, such issues seem to be problematic for doctors across different health care systems.

With respect to Karasek’s demand–control model, although we did not test this model, in the present study perceived job demands were associated with life satisfaction, but not autonomy. This may be because Norwegian doctors overall have less professional autonomy within the public system than, for example, US doctors in a private system (268). Nevertheless, we included only a few variables resembling Karasek’s measures, which is a study limitation.

Colleague support was directly linked to doctors’ well-being in this study, which is
consistent with previous research (158, 269). Next to work–home stress ($t = –8.7$), colleague support ($t = 7.0$) had the highest effect size of the three significant work-related factors in our study. Support from peers and a good working environment may buffer some of the stressors related to doctors’ work (270). A study of Scottish health care professionals in oncology showed that co-worker support was associated with perceptions of reduced effort, greater reward and increased satisfaction (271). Several studies in Europe and the USA have emphasized the importance of colleague support and better teamwork and communication, in both practice venues and hospitals, for physician satisfaction, findings that are in keeping with ours (272, 273). A major strength of our study is that we controlled for other types of social support, including both perceived and structural (i.e., having a partner) support, as well as personality factors. Thus, the importance of colleague support at work should be emphasized with respect to interventions fostering life satisfaction among doctors.

5.8 Other individual predictors or factors linked to life satisfaction

As expected, we identified some individual factors already known to be associated with life satisfaction, such as being married/cohabiting, perceived social support, physical activity, the personality trait of neuroticism (a true predictor) and life events. In addition, we also identified some new individual factors linked to life satisfaction: reality weakness personality trait (a predictor), female gender and drinking behaviours (i.e., hazardous drinking and drinking to cope with tension). Drinking behaviours will be discussed separately.

**Reality weakness personality trait** was an independent predictor of poor life satisfaction even after controlling for other personality traits and other factors. This is a novel finding. This personality trait is related to ‘the lack of ability to handle inner impulse and to resist outer stimulations’ (190, 274). Individuals with this personality trait have identity problems and are suspicious of others. They are on the borderline between reality and fantasy, and have been associated with paranoid, borderline and schizotypal personality disorders. In several previous NORDOC studies, our group investigated the predictive validity of reality weakness as a personality trait among Norwegian doctors, with regard to their need for mental health treatment during medical school (201), aggravation of suicidal ideation (45), lack of help-seeking for mental health problems (275) and severe depressive symptoms (276). We also know that individuals with personality problems, and even distortions in their self-perception or identity, also have difficult relationships with others. This may lead to reduced well-being
and life satisfaction. In another Norwegian cohort study, the personality trait of reality weakness was also linked to perceived job stress and symptoms of depression and anxiety among junior doctors during their postgraduate training (195). In sum, the reality weakness personality trait is an important predictor of depression and aggravation of suicidal ideation during the initial postgraduate years; in the present study, this trait was shown to reduce life satisfaction over a 15-year follow-up period. However, because the reliability of this variable in our sample was questionable (Cronbach’s alpha = 0.62), this finding should be replicated using a stronger measure (e.g., with a Likert scale).

In our unadjusted model, women reported similar levels of life satisfaction compared with those reported by men. In our fully adjusted model, female gender was associated with higher levels of life satisfaction. This is consistent with a recent Norwegian prospective longitudinal study of the general population that showed that females scored higher on life satisfaction than did men (277). In our study, female gender became significant in Model 1 when the personality trait of neuroticism was entered. This means that the effect of gender had been suppressed by the strong effect of neuroticism on life satisfaction and the relatively high level of neuroticism among female students (mean = 4.01, SD = 2.16) compared with male students (mean = 2.93, SD = 2.15). On the other hand, several US studies of doctors have shown unadjusted lower levels of well-being among women than among their male colleagues (120, 149); our finding of similar satisfaction levels between the genders may be due to the relatively egalitarian nature of Scandinavian society.

**Perceived social support** had an independent, positive effect on life satisfaction among doctors, which is consistent with other cross-sectional studies of doctors (120, 148, 278) and prospective studies of the general population (169, 279). However, to our knowledge, the present work is the first prospective longitudinal study showing social support to be important throughout the first 15 years of doctors’ careers. A previous NORDOC study found that about half the variance in life satisfaction during the early career period was explained by both structural support (i.e., having a partner) and perceived social support (20). Thus, appreciation by, for example, close friends and family members may serve as a form of structural, practical or emotional support, which may increase quality of life among doctors.
Life events had a direct impact on life satisfaction among Norwegian doctors. This is not an original finding but is in keeping with previous research (169, 280-282). Our post hoc univariate analyses showed that life events predicting lower life satisfaction included: a serious disease/accident/hospital admission; divorce/separation/broken relationship; other difficulties among close family; serious financial problems; serious problems with your residence/dwelling; partner being unemployed/granted leave; you, or someone in the close family, having been involved in a serious violation of the law; problems with your partner; and other serious events (self-specified). Such life events were associated with lower life satisfaction among doctors.

5.9 Predictors of a decline in life satisfaction

A high level of the neuroticism personality trait (self-criticism/low self-esteem) and a low level of colleague support were the only factors predicting a decrease in life satisfaction, further emphasizing the importance of these variables. We successfully validated this finding by including both the stable group and the group with increasing life satisfaction as a reference. Previously, we found that personality trait (low conscientiousness) predicted an increase in life satisfaction during the first postgraduate year (20), but we are unaware of any other prospective studies on the role of personality in doctors’ positive psychology.

Prospective studies among doctors have shown the personality trait of neuroticism to be a risk factor for negative psychology and mental health, including depressive symptoms (139, 276, 283, 284), suicidal ideation (45, 207), stress (170, 195) and burn-out (170-172, 206, 285). In conclusion, these findings indicate that the personality trait of neuroticism may be a risk factor not only for depression, but also for reduced life satisfaction in general. Studies in other populations have also found that neuroticism predicts subjective well-being (286-288).

This is the first prospective longitudinal study to demonstrate that low colleague support predicts a decline in doctors’ life satisfaction and that it is a true predictor. A recent NORDOC study showed that reduced colleague support predicts emotional exhaustion five years post-graduation, particularly among men (210). In another qualitative study by our group, specialist registrars reported that in a doctor’s work ‘one cannot be ill’ (259). In that study, having colleagues absent from work meant more hectic days and less time for professional development for those left working, and that unstaffed work/night call shifts meant shorter rest periods. Absentee specialist registrars added additional work and also led to unpredictability for senior consultants. In addition, the absence of individual doctors was
frequently perceived as inadequate loyalty to their colleagues. In sum, many aspects of a doctor’s work life are very demanding. Most of the time doctors work alone and have significant responsibilities; therefore, support from their colleagues is extremely important to their life satisfaction.

5.10 Drinking behaviours independently linked to life satisfaction

It is a novel finding that among doctors, both drinking to cope with tension and our measure of hazardous drinking (AUDIT-9) were independently associated with poor life satisfaction after controlling for other individual- and work-related factors. Drinking to cope with tension is a type of avoidance behaviour that may theoretically reduce quality of life during one’s career, especially a career that requires active coping with frequently demanding patient work. Another long-term prospective NORDOC study found that drinking to cope in medical school predicted reduced perceived mastery of clinical work 10 years after leaving medical school (289). This type of drinking behaviour should be prevented and healthier coping strategies should be adapted. Our finding with regard to hazardous drinking and life satisfaction is consistent with a recent large cross-sectional survey of young adults attending European universities, which showed that harmful or hazardous alcohol consumption was related to a massive decrease in quality of life (290). One limitation of the present study is that we cannot know whether the relationship between life satisfaction and hazardous drinking is direct or reciprocal. We do not know which comes first: drinking behaviour or reduced life satisfaction. This issue needs to be further clarified using a prospective study.

5.11 Gender differences in predictors of hazardous drinking

A significant interaction was found between mental distress and gender, with mental distress a more important predictor among men than among women. There are several potential reasons for this. First, drinking is traditionally a male behaviour, thus men’s threshold for ameliorating emotional distress by drinking to intoxication may be lower; men may also avoid seeking professional help (291). Women are more likely to see a doctor and use prescribed tranquillizers or anti-depressants (292, 293). Lack of help-seeking for mental health problems in doctors is an issue that has also been under investigation for many years (248), and we know that among doctors in Norway, more women seek counselling for burn-out compared with men (206).

There was a significant interaction between gender and life events, indicating that negative life events are more important predictors among men compared with women. One
reason for this may be that men tend to use externalizing behaviour (e.g., physical aggression, substance use, drinking) more often than women when they experience life events (294). Furthermore, we know that men are more prone to suffer after a divorce, and that they encounter serious legal issues and financial problems more often compared with women. In conclusion, doctors should be aware that experiencing mental distress and life stress may lead to hazardous drinking and they should be encouraged to seek professional help when it is necessary.

5.12 Gender differences in predictors of life satisfaction

There was a significant interaction between gender and use of alcohol to cope with tension, indicating that among men who used alcohol as a coping strategy, there was a stronger reduction in life satisfaction compared with women using a similar strategy. This may be explained by the fact that drinking is a male behaviour, insofar as men drink more often and in greater quantities, and that men and women tend to use different means of coping with stress. In general, men report more drinking-related problems, as well as more negative consequences to their well-being, than do women (70).

There was a significant interaction between neuroticism and gender, indicating that the negative impact of neuroticism on life satisfaction was stronger among men compared with women. We can only speculate as to the reason for this. It may be that vulnerability is more difficult for male doctors to deal with over the years than it is for their female colleagues. Although gender roles are becoming more equal, lingering societal expectations may suggest that male doctors should be stronger and more independent than their female colleagues. Doctors’ work often demands prompt action, at least in emergency cases, and is a requirement that may be more difficult for male doctors high in neuroticism and uncertainty. We know from another study that female doctors receive social support from their colleagues more readily, for instance to buffer work–home stress (114).

Furthermore, there was a statistically significant interaction between perceived job demands and gender, in that the negative impact of perceived job demands on life satisfaction was stronger among men than women. There may be several reasons for this finding. First, treating demanding patients may be easier for women, who have been shown in a meta-analysis of 26 studies to be better communicators compared with men (295). Female doctors also tend to be more patient-centred in their communication and have longer visits with their patients than do male doctors. Second, women choose part-time work more often and may
give up a stressful medical career because of family demands; thus, men may simply spend more time in stressful job settings and work longer hours (114). Third, women doctors receive more support from their colleagues (114).

5.13 Cohort effects

There was a cohort effect on the association between work–home stress and life satisfaction, indicating a significantly stronger association among the younger cohort of doctors (the Medical Student Cohort, which graduated in 1999). Additional analyses showed that the Young Doctor Cohort reported the highest levels on all perceived job stress dimensions, although the strongest effect of work–home stress on life satisfaction was found among the youngest cohort (the Medical Student Cohort) (Pearson’s r = -0.34 vs. -0.26, respectively). This was an unexpected finding. When we controlled for both number of children and hours at work/week, we found no significant difference between the cohorts. Previous longitudinal studies using the Young Doctor Cohort have shown that work–home stress varies over the medical career, with an increase in the first postgraduate years and a subsequent reduction 10–15 years after graduation (114, 210). With respect to cohort differences in NORDOC, a recent study (Hertzberg et al., resubmitted) has shown that the youngest cohort (the Medical Student Cohort) experienced the least work–home stress in a fully adjusted predictor model. This may be due to improved social welfare for families with children over the last decade in Norway (e.g., full kindergarten coverage), which would have preferentially benefitted the youngest cohort; thus, the protective effect of such measures on work–home stress and subsequent life satisfaction may be more important in the younger cohort. There may also be more variation between individual families with respect to family support, since typically both parents must work due to considerable accommodation and living expenses.

5.14 Implications for medical school and clinical practice

Hazardous drinking among doctors may result in their having serious health problems, as well as impacting on their work capacity and medical practice. We know from previous research that hazardous drinking over time increases the risk of alcohol dependency, although our study reports that among those who drank hazardously in medical school, only 23% (n = 18) continued this pattern 10 years post-graduation. Thus, three out of four doctors discontinued such behaviour after leaving medical school, likely because they established a family, starting working, etc. We know that drinking is very common at social gatherings among medical students. Recent preliminary data from Trondheim and Bergen (Reidar
Tyssen, personal communication) show more hazardous drinking among medical students with lower levels of anxiety and depressive symptoms (SCL-5), whereas students who did not drink hazarously reported more of these symptoms. This may have to do with social exclusion and stigma related to abstinence (e.g., among those affiliated with a religion). Studying medicine obviously does not in itself prevent harmful use of alcohol. In a Danish study, 50% of doctors reported having limited knowledge about substance use disorders, or being dissatisfied with the knowledge they acquired during their medical education about substance use disorders (62). Therefore, in developing preventive efforts, we must examine how alcohol-related problems are taught in the medical curricula, and the extent to which student organizations may contribute to students’ use of alcohol.

Another important lesson from this study is that using alcohol to cope with tension may be a problem for medical students and doctors. This manner of coping with stress should therefore be highlighted when medical students and young doctors are taught about general strategies for coping with stress or distress. Medical schools should also promote knowledge about substance use disorders, while simultaneously encouraging students to adopt healthier coping strategies (i.e., rather than using avoidance strategies/self-medication with alcohol to cope with tension). In addition, vulnerable personality traits and mental distress should be addressed among medical students. In particular, those with vulnerable personality traits and those experiencing negative life events should be encouraged to pay attention to possible protective factors or strategies against mental distress/illness. Furthermore, that drinking may reduce impulse control should be a focus when teaching students about the consequences of drinking. Medical students and doctors should not delay seeking help to combat mental distress/illness and problematic drinking behaviours. Intervening as early as possible is important, not only for the health of the doctor, but for their patients.

Doctors’ life satisfaction is essential to their delivery of high-quality health care. Doctors are the backbone of a nation’s health care system. Increased awareness of the importance of doctors’ life satisfaction should be emphasized at the workplace, both among doctors themselves and their employers. Preventive measures should be taken at the organizational level, including dedicating sufficient resources to reduce doctors’ workloads when illness occurs, to help prevent them from becoming dissatisfied and reducing their well-being. In Norway, as in other Scandinavian countries, doctors’ working hours are highly regulated yet perceived job demands, work–home interference and lack of colleague support are factors clearly associated with their poor life satisfaction. As noted, Eckleberry-Hunt et al.
showed that rather than with the number of working hours/week, the perception of a heavy workload is associated with lower physician wellness (17). Interventions at the organizational level aimed at improving the working environment (e.g., reducing workload, promoting flexibility and work-life integration) may lead to enduring change and reduced dissatisfaction among doctors. US studies have found that better organized teamwork at practice venues can also increase physician well-being (273). Increasing the number of medical assistants per doctor may even be financially advantageous, since this practice increases patient turnover, productivity and provider satisfaction (296). Balancing one’s professional and personal life, in addition to promoting colleague and co-worker support in the workplace, should also be emphasized.

Individual factors such as being married/cohabiting and enjoying favourable social support should also be on the agenda when teaching students and young doctors about their future well-being. Healthier lifestyle habits should be adopted and those with vulnerable personalities should seek help.

5.15 Implications for future research

We did not find any associations between work-related factors and hazardous drinking. Future research should nevertheless investigate the impact of work-related predictors of hazardous drinking among doctors in regions other than Scandinavia, where work–life balance is less regulated and working hours are longer. Future studies should also examine whether drinking to cope with tension is a mediator or intermediate factor between work-related stress and hazardous drinking.

Despite this being a prospective longitudinal study, we cannot deduce from these results that mental distress causes hazardous drinking or that hazardous drinking is caused by mental distress. Therefore, further research is also needed to address the existence and direction of potential causal relations. Given the opportunity to study hazardous drinking and mental distress among doctors again, I would perform ‘lagged analyses’ to determine the direction of causality (if causality exists). Temporality is an essential criterion with respect to possible causation and identification of risk factors (216). Our study results suggest new research questions, such as which characteristics of individuals lead to stable, high life satisfaction over time. Furthermore, with an increasing proportion of female doctors in the medical profession (as in other Scandinavian countries), it will be important to assess the impacts of this shift on the community of practising doctors as a whole.
5.16 Conclusions

Hazardous drinking
- There are several individual-based, independent risk factors present during medical school that predict later hazardous drinking.
- Drinking to cope with tension during medical school and low levels of the personality trait of conscientiousness are independent predictors of hazardous drinking.
- Doctors’ work-related stress is not associated with hazardous drinking; however, co-occurring life events and mental distress are independently associated with hazardous drinking.

Life satisfaction
- Both drinking to cope with tension and hazardous drinking are independently and negatively associated with life satisfaction.
- With regard to work and life satisfaction, work–home stress and perceived job demands are both negatively associated with life satisfaction, whereas colleague support is positively linked to life satisfaction after controlling for other factors.
- The personality trait of reality weakness is an independent predictor of poor life satisfaction.
- Both high levels of the neuroticism trait and low colleague support independently predict decreasing life satisfaction over time.

Major findings with regard to doctors’ work
- This long-term, longitudinal study of Norwegian doctors failed to find a link between work and hazardous drinking, whereas work-related factors were linked to life satisfaction, even after controlling for individual factors.
REFERENCES


121. Yamey G, Wilkes M. Promoting wellbeing among doctors: We should move away from a disease model and focus on positive functioning. BMJ. 2001;322:252.


147. West CP, Shanafelt TD, Kolars JC. Quality of life, burnout, educational debt, and medical knowledge among internal medicine residents. JAMA. 2011;306:952-60.


176. Wiers-Jenssen J. Rekruttering til medisinstudiet: en studie av et årskull medisinstudenter, med vekt på sosial bakgrunn, poengsamling, studiemotiver og ambisjonsnivå/Recruitment to medical study. A study of a class of medical students with emphasis on social backgrounds, collection of study points, study motives and ambition level (in Norwegian) [Master]. Oslo: University of Oslo; 1994.


286. Vittersø J, Nilsen F. The conceptual and relational structure of subjective well-being, neuroticism, and extraversion: once again, neuroticism is the important predictor of happiness. Soc Indic Res. 2002;57:89-118.


