

RESEARCH

Open Access



Using causal effect estimation to evaluate occupational stress factors during a weekend on-call shift among general practitioners in training

Géraldine Deberdt^{1*}, Paloma Rabaey², Sam Van Damme², Laure-Ann Quenon¹, Peter Decat¹ and Thomas Demeester²

Abstract

Background In recent years, our understanding of the various stress factors affecting the mental well-being of medical trainees during on-call work has grown. Although we have developed a general understanding, we still lack insight into the underlying interrelationships and causal effects of these factors, necessary to address and prioritize the mental health needs of young doctors.

Objectives We aimed to measure the causal effects of selected stress factors related to on-call work contributing to occupational stress among General Practitioners (GPs) in training.

Methods We identified key stress factors through a literature review and group discussion. Furthermore, we built a Directed Acyclic Graph (DAG) for each identified factor to draw valid causal inferences based on data collected from an observational survey.

Results Our analysis indicated that GPs in training experience interruptions to leisure time, fear of potential work-related errors, and anticipatory and vigilant behaviour towards potential patient complaints as key contributors to elevated stress levels during on-call work.

Conclusion Our study revealed the importance of psychological or emotional stressors during an on-call shift encountered by GPs in training, encouraging preventive measures.

Trial registration Not applicable.

Keywords Causal effect estimation, On-call shift, General Practitioner in training, Directed Acyclic Graph

Peter Decat and Thomas Demeester shared last authorship.

*Correspondence:

Géraldine Deberdt

geraldine.deberdt@ugent.be

¹Department of Public Health and Primary Care, Ghent University, Corneel Heymanslaan 10, Ghent 9000, Belgium

²IDLab, Department of Information Technology (INTEC), Ghent University-imec, Technologiepark-Zwijnaarde 126, Ghent 9052, Belgium



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Background

Physician well-being is a growing concern within the medical profession. In recent years, the incidence of mental health problems among physicians has been increasing significantly [1–3]. Particularly, medical trainees appear to experience severe work-related stress and burnout. Dyrbye et al. [4] recognize the impact of work-related stress on trainees' professional development, the ideation of suicidal intentions and on the safety of patient care. Although physician well-being is frequently discussed, research into the causes and risk factors of occupational stress is not given enough priority [5]. Considering its widespread prevalence and serious repercussions, preventive measures are needed to tackle this problem.

Research reveals that out-of-hours services significantly contribute to occupational stress among General Practitioners (GPs) in training [6, 7]. For instance, related work states that on-call duty in unfavorable working conditions is associated with medical errors, decreased well-being, fatigue, and burnout [6–8]. Nevertheless, on-call services remain necessary and inevitable for maintaining continuity of care. Therefore, it is essential to identify the various factors that increase occupational stress during out-of-hours services among GPs in training to gain a better and more detailed understanding of the drivers of stress in this context, which can then inform effective prevention strategies.

Previous research in secondary care has identified several characteristics of on-call shifts influencing occupational stress. In addition, demographic factors such as being female have been associated with mental fatigue and burnout, while younger age has been linked to lower job control and higher levels of experienced job strain [5, 7, 9]. Underlying mechanisms, such as physicians' ruminations about interruptions to family or leisure time, concerns about missing a call, pre-planning for future calls and the inability to make plans, have been clarified [10]. On the other hand, organizational justice [7], the presence of a recovery period post-call [11, 12], social support [5], a healthy team climate [7, 11] and a supporting supervisor available [7, 13] proved to be effective job resources balancing the experienced on-call demands and stress.

Although we have a general understanding of the various factors contributing to occupational stress on-call in widespread medical specialties, we lack insight into the key stress factors, their interrelations and individual causal effect on occupational stress, insights that are essential for developing effective preventive strategies. Within primary healthcare, particularly in GPs in training who frequently undertake on-call shifts, there remains a lack in research addressing the unique needs of young doctors. For this reason, we set out to measure the

causal effects of stress factors of interest, as determined in a literature review and group discussion. We conduct an observational survey among GPs in training in order to obtain data on the retained stress factors in relation to occupational stress during a weekend on-call shift. Furthermore, we use a Directed Acyclic Graph (DAG) for each identified stress factor to draw valid causal inferences based on observational data.

Methods

Identification of stress factors

We identified potential stress factors contributing to occupational stress during on-call shifts, based on a literature review and a group discussion with experience experts. A literature review was conducted by consulting Google Scholar and the PubMed database for English-language papers published since 2004 to comprehend the key stressors. A detailed search strategy is provided in [Appendix 1](#). The identified stress factors were then presented to a group of 11 GPs in training with experience in performing on-call shifts, to gather clinical experiences that validate the literature findings and, if necessary, complement them with other relevant factors that may have been missed previously through literature search alone. The buildup of the group discussion was based on recent recommendations from Mortelmans D. and Steenhuyzen S. et al [14, 15] and is provided in [Appendix 2](#). Eventually, 22 factors were retained as eligible stress factors during an on-call shift and consequently included in the observational survey.

Observational survey

Study protocol and measurements

We set out to conduct an observational study to obtain data on these retained stress factors. As shown in [Fig. 1](#), GPs in training were asked to complete two questionnaires during a scheduled weekend on-call shift. The first questionnaire, available in [Appendix 3](#), evaluated the properties and characteristics of the on-call shift, measuring the 22 factors that were previously identified through literature review and group discussion. The second questionnaire assessed the level of occupational stress respondents experienced during this shift using the validated short form 6a 'TBI-CareQOL Caregiver Strain,' as detailed in [Appendix 4](#) [16]. The questionnaire had to be completed as soon as possible, within a maximum timeframe of three days post-call.

Study setting and participants

The survey was conducted from October 2023 to February 2024, collecting data from different GP cooperatives¹

¹A GP cooperative is a large-scale organization of about 15 to more than 250 GPs. GPs rotate shifts to provide out-of-hours care for the combined

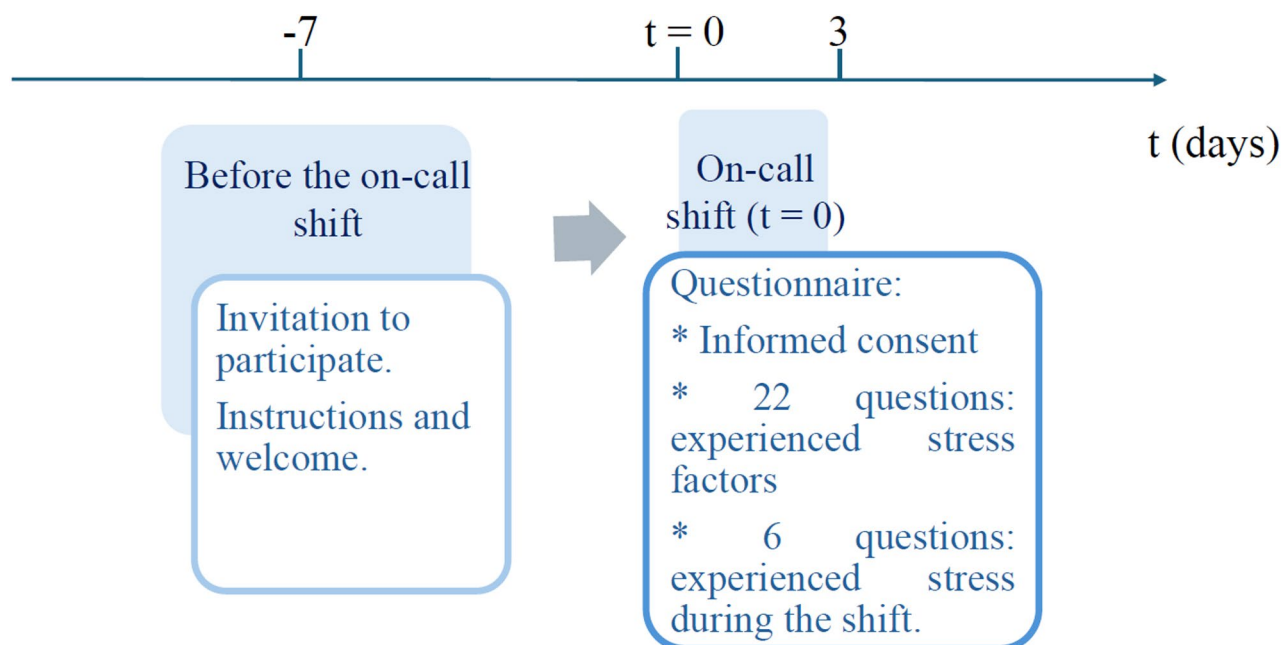


Fig. 1 Overview of study protocol

in Flanders (Belgium), mainly in the cities of Gent, Brugge, Antwerpen and Vlaams-Brabant. Our inclusion criteria encompassed all GP trainees (first, second, or third year) scheduled for a weekend on-call shift in the aforementioned regions. Specifically, participants had to be healthy, mentally competent individuals capable of providing consent for the study. Only trainees with an atypical training trajectory, such as part-time employment, were excluded.

Causal effect estimation

Based on the obtained observational data, we aim to uncover the causal effect of 9 stress factors of interest. These factors were selected based on (i) the frequency with which they were mentioned in literature, (ii) the frequency with which they were brought up by participants in group discussion, as well whether they were discussed at length and whether participants agreed on their importance, and (iii) expert input. To this end, we built a separate DAG for each factor. DAGs offer an innovative approach to identify confounding variables and estimate causal effects in applied health research [18]. Although the importance of DAGs in causal research is widely recognized, the key guidelines and recommendations providing specific steps for DAG development remain limited and inconsistent. We used a strategy proposed by

Poppe et al. consisting of six steps to build causal DAGs [19].

The constructed DAGs and their descriptions were presented to a domain expert, a GP who was familiar with the framework of DAGs but otherwise not involved in this study. He reviewed the DAGs critically and proposed some adaptations to cluster nodes, to add ancestors and refine the terminology used. Furthermore, the factors and interrelations that remained subject to discussion were evaluated through a brief questionnaire by a group of 14 GPs in training who were not otherwise involved in the study. Feedback from the domain expert, as well as from the 14 GPs in training were integrated, finalizing the DAGs. We used *DAGitty version 3.1*² to create editable DAGs and provide coding details, accessible to readers at <https://dagitty.net/mabyRKJGP>. Next, we used our DAGs to identify minimally sufficient adjustment sets needed to estimate a causal effect of a factor of interest X on the outcome Y (occupational stress) without bias from confounders or improper adjustment. We used the following strategy to test the causal effect of each variable of interest.

1. Designate the variable for which we want to estimate a causal effect as the exposure variable X . On-call stress remains the outcome variable Y .
2. Identify a minimal adjustment set. This is a set of variables that we can adjust for in order to block all paths from exposure X (candidate stress factor) to

patient population of all participating GPs. Such GP cooperatives are typically observed in countries like Belgium, the Netherlands, Germany, and other parts of Europe [17].

²<https://dagitty.net/dags.html>

outcome Y (stress). We refer to these variables as C (confounders).

3. Fit a linear regression model for the outcome Y , regressing variables X and C .
4. The p -value of X according to the Analysis of Variance (ANOVA) test tells us whether there is a significant effect of X on Y when controlling for C . If so, the coefficients of the various categories of X in the regression model inform us by how many units stress is increased or decreased on a total of 30 units in the ‘TBI-CareQOL Caregiver Strain’ scale per category, as a result of X .

Results

First, we present the obtained demographic data and outcomes from the observational survey on the 22 initially retained stress factors. Subsequently, in [Identified stress factors](#) section, we guide through the nine key stress factors during an on-call shift which we selected as most relevant based on (i) the frequency with which they were mentioned in literature, (ii) the frequency with which they were brought up by participants in group discussion, as well whether they were discussed at length and whether participants agreed on their importance, and (iii) expert input. Finally, in [Causal analysis](#) section, we built a DAG for each of these 9 stress factors of interest to enable the inference of valid causal conclusions.

Demographic features and survey outcomes

A total of 76 participants answered our email invitation to participate in the study, with 69 of these participants eventually completing the observational study. We excluded five participants in total: two because they completed the questionnaire during a weekday shift instead of a weekend shift, and another three who misinterpreted the questionnaire. 64 respondents were included in the final analysis, of which [Table 1](#) displays the demographic characteristics. Note that ‘work environment’ is limited to a single category of the weekend on-call center setting

Table 1 Demographic features of respondents

Characteristic		Number	Percentage, %
Gender	Male	11	17.2
	Female	53	82.8
	Non-binary	0	0
Experience	1st or 2nd year	34	53.1
	3rd year	30	46.9
Work environment	On-call center setting	64	100
Area of GP cooperative	Gent	26	40.6
	Brugge	10	15.6
	Antwerpen	15	23.4
	Vlaams-Brabant	13	20.3

at GP cooperatives. Restricting the data instances to a single type of work environment implies controlling for that variable by design, indicated further in the DAGs as a white node. The assessed stress factors, response categories provided to the respondents and distribution of the results are illustrated in [Table 2](#). The factors identified as the most prominent stressors contributing to occupational stress in an on-call setting, and used in the causal analysis, are highlighted in bold.

Identified stress factors

In total, we selected nine factors as the most prominent stressors contributing to occupational stress in an on-call setting based on literature review and group discussion: interruptions to family or leisure time; anticipating action; potential errors; hourly caseload; post-call recovery; sex; experience; unfair questions and demanding behaviour. Of these, five were initially recognized in literature and afterwards supported by the group discussion (interruptions to family or leisure time; anticipating action; potential errors; experience; post-call recovery). Two additional factors appeared exclusively in literature (sex; hourly caseload) but were not confirmed in the discussion. The last two factors were noticed during the group discussion alone (unfair questions; demanding behaviour). We provide a description and further details of the selected stress factors, illustrated by quotes (translated from Dutch to English) and literature references in [Appendix 5](#).

Anticipatory behaviour prior to future calls or patient visits, the experienced interruptions to leisure time due to an on-call shift and the fear of possibly making medical errors emerged initially as key contributors to elevated stress levels in literature and overlapped with the content of the group discussion [6, 10, 20]. For instance, respondents expressed constant vigilance over their phones and fear of missing a call. During breaks between patients, they struggle to relax and remain alert, making sleep difficult. Additionally, they anticipate scheduled patient visits and preoccupy themselves considering the potential differential diagnoses beforehand.

‘It might sound odd, but your phone can ring at any moment. This leads to a lot of practical issues, for instance questions such as “Should I go to the bathroom now?,” “When should I eat?,” “When should I change clothes and brush my teeth?” Being constantly aware and vigilant burdens me, it even triggers anxiety.’

Also, the physician’s gender and experience proved important. As an example, Vincent et al. found that female on-call workers appeared to exhibit more anticipatory behaviour, such as thinking about the likelihood of

Table 2 Responses to questionnaire items

Stress factor of interest	Assessed categories	Number	%
Hourly caseload	Busy	36	56.3
	Average	16	25.0
	Calm	12	18.8
Post-call recovery	I have all day to recover	33	51.6
	I have to work the whole day as usual	25	39.1
	I have some hours recovery time	6	9.4
Interruptions to leisure time	Experiencing some interruptions	34	53.1
	Experiencing many interruptions	26	40.6
	Experiencing few interruptions	4	6.3
Unfair questions	Rather unjustified registrations	44	68.8
	Rather valid registrations	20	31.3
Demanding behaviour	Little patients were demanding	49	76.6
	Some patients were demanding	14	21.9
	Many patients were demanding	1	1.6
Potential errors	Rather worried about potential errors	27	42.2
	Sometimes worried about potential errors	25	39.1
	Little worried about potential errors	12	18.8
Anticipating action	Highly recognizable	39	61.0
	Recognizable to a limited extent	25	39.1
Medical support	Well supported	44	68.8
	Average	13	20.3
	Limited support	7	11.0
Interruptions to social activities	Experiencing some interruptions	38	59.4
	Experiencing many interruptions	20	31.3
	Experiencing few interruptions	6	9.4
Unknown medical records	Accessible but complicated	41	64.1
	Easy access	18	28.1
	No access	5	7.8
A driver	The driver proved to be very useful	29	45.3
	The driver proved useful	17	26.6
	Neutral	9	14.1
	No driver present	6	9.4
	Driver present but not very useful	3	4.7
A secretary's office	The secretary office proved useful	30	46.9
	The secretary office proved to be very useful	26	40.6
	Neutral	8	12.5
Unsafe situations	No unsafe situations encountered	61	95.3
	I have felt unsafe to a limited extent	3	4.7
Practical barriers	No problems experienced	28	43.8
	Experiencing problems to a limited extent	21	32.8
	Experiencing some problems	15	23.4
A medical triage	Medical triage proved to be of little use	55	86.0
	The medical triage proved useful	9	14.1
Tolerance level	Calm and controlled towards patients	32	50.0
	Neutral	21	32.8
	Irritable and not very tolerant of patients	11	17.2
Hospital resources	Average	27	42.2
	Well supported	24	37.5
	Limited support	13	20.3
Socio-economic state (SES)	Middle/high SES	46	71.9
	Low SES	18	28.1

We depict the assessed on-call stress factors in the questionnaire (first column, the corresponding questions used in the survey are available in [Appendix 3](#)), with the possible response options provided to the respondents alongside (second column). The distribution of answers is shown in the third and fourth column. The response options used in the questionnaire were administered in Dutch, the respondents' native language, and had been translated here from Dutch to English. The factors identified as the most prominent stressors contributing to occupational stress in an on-call setting, and used in the causal analysis, are highlighted in bold

being called and reporting thoughts about pre-planning their visits [10]. Notably, some participants mentioned that these ruminations evolved depending on the number of years of experience.

‘The stress before a shift does get better with experience. I still remember my first shift, “hallelujah”, whereas now I believe I have more knowledge. You trust yourself a lot more. I know what I need to do, and if I do not trust it, then I’ll triage to the emergency department. If I was a first-year trainee General Practitioner, I would have panicked more.’

Furthermore, literature highlights the importance of a recovery period after shifts as a valuable resource for balancing the demands of on-call work [11, 12]. Respondents confirmed that a post-call recovery period would provide relief, however many noted the strong taboo still surrounding this topic.

Table 3 The table presents the results from the causal analysis for relevant exposure variables, listed in the first column

Exposure X (Reference)	Model	p-value (ANOVA)	Coefficients	95 % CI
Interruptions to leisure time (Ref: Experienced some interruptions)	On-call stress ~ Interruptions to leisure time + Anticipating action + Experience + Hourly caseload + Post-call recovery + Sex + Tolerance level	0.0026	Experienced many interruptions = 2.4959 Experienced little interruptions = -1.1616	[0.799, 4.193] [-4.800, 2.477]
Anticipating action (Ref: Only slightly recognizable)	On-call stress ~ Anticipating action + a Driver + a Secretary's office + Experience + Hourly caseload + Medical support + Potential errors + Sex	0.0002	Easily recognizable = 2.9372	[0.933, 4.942]
Potential errors (Ref: Rather worried)	On-call stress ~ Experience + Hospital resources + Medical support + Sex + Tolerance level + Unknown medical records	0.0295	Slightly worried = -2.8845 Occasionally worried = -1.3633	[-5.730, -0.039] [-3.663, 0.936]

The second column indicates the model used, which is a regression model of on-call stress as a function of exposure and required confounders. The *p*-value for each exposure is reported to indicate the presence of significant differences in on-call stress among the different categories of the exposure. The fourth column shows the difference in mean values of the indicated categories of the exposure variable with respect to the reference category (column Coefficients; reference mentioned in parentheses). These differences are measured in units of stress, with a total of 30 units on the ‘TBI-CareQOL Caregiver Strain’ scale. The right-most column provides the corresponding 95% confidence interval

‘In my opinion, recovery time would be a relief! Sometimes, I’ve had a busy night, which affects my sleep for several days afterwards. It’s indeed unfortunate that there exists such a taboo, especially considering that we’ve already worked continuously for 24 hours. Moreover, recovery time is actually stipulated in our contract.’

In addition, numerous studies highlighted the association between an increasing hourly caseload on-call, defined as the number of patients or cases presenting each hour during the on-call shift, and mental fatigue or occupational stress [5, 9, 21]. Through the group discussion, we identified the importance of two final stress factors. On the one hand, participants highlighted that some patients exhibit demanding behaviour. On the other hand, they noted unjustified patients’ complaints, citing examples such as earwax removal or administrative requests related to disability certificates.

‘While accessibility is important, General Practitioners are not to be taken advantage of.’

Causal analysis

Results from the causal analysis are depicted in Table 3. The calculations were performed based on the constructed DAGs created for each stress factor of interest, as shown in Fig. 2. The exposure *X* is represented in the first column, alongside the corresponding regression model built with the confounders we have to account for. Subsequently, we represent the *p*-value assigned to the exposure in the ANOVA test. The coefficients corresponding to the various categories of each exposure variable and their 95%-Confidence Intervals (CIs) are provided in the last two columns. The coefficients are measured in units of stress based on the ‘TBI-CareQOL Caregiver Strain’ scale, with a maximum of 30 units. Note that each category is plotted against a reference or baseline category indicated in parentheses. Only the stress factors ‘interruptions to leisure time’, ‘anticipating action’ and ‘potential errors’ have a significant effect on the outcome according to our analysis. For these variables, we examine which categories are associated with elevated stress levels. For ‘post-call recovery’, ‘hourly caseload’, ‘sex’, ‘experience’, ‘unfair questions’ and ‘demanding behaviour’ the *p*-values were respectively 0.40, 0.54, 0.86, 0.52, 0.99, and 0.64, indicating no significant effect on stress.

Resting on the assumption that our DAG is correct, we report the following findings regarding the causal effects on stress:

1. Doctors experiencing many interruptions in leisure time due to an on-call shift have a stress level that

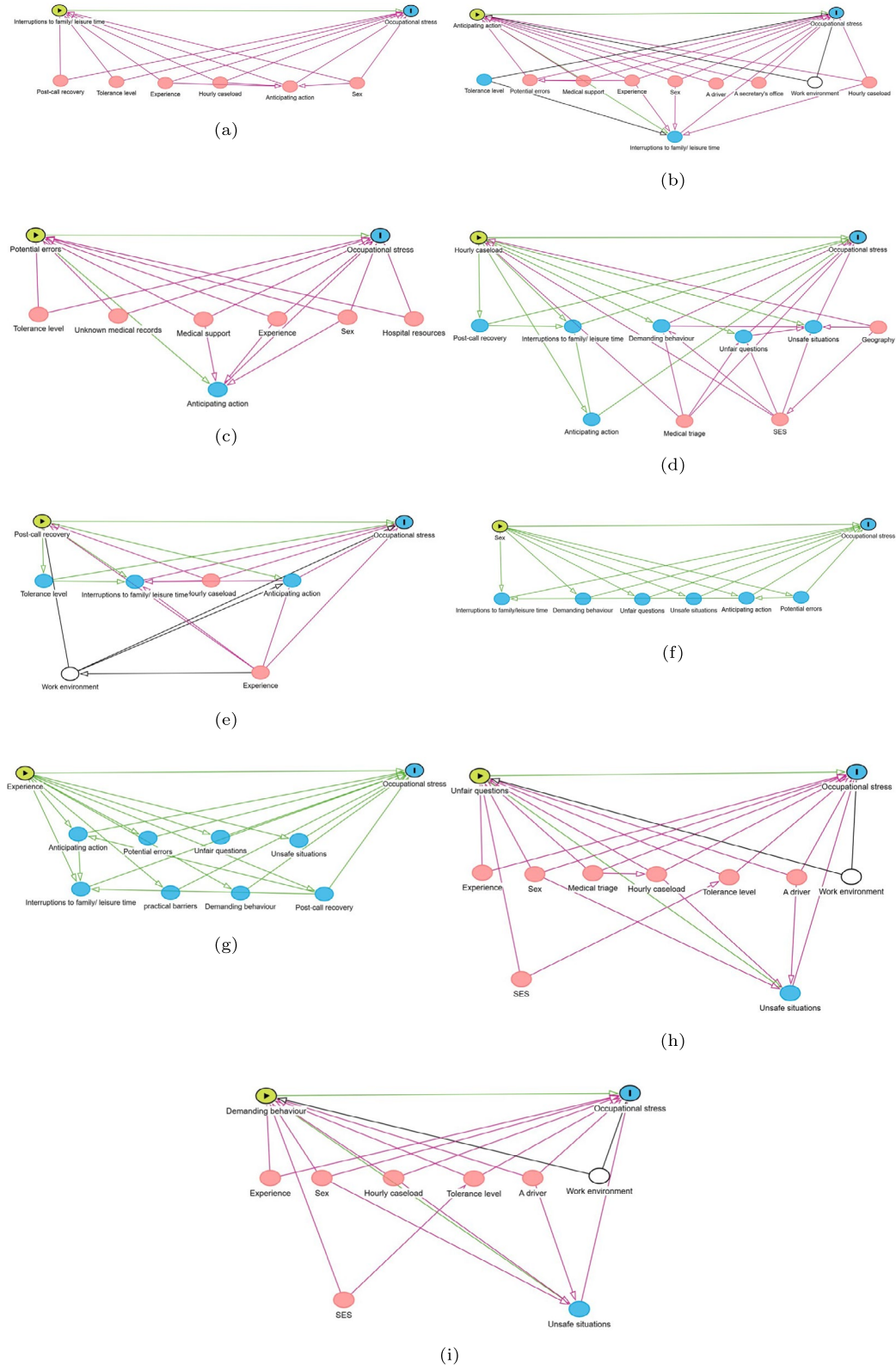


Fig. 2 Constructed DAGs representing the causal structure for the association between a stress factor of interest and occupational stress on-call. The light green node represents the exposure (i.e. stress factor of interest), the surrounded blue node represents the outcome (occupational stress), other blue nodes represent ancestors of the outcome and the white node represents an adjusted factor 'work environment' controlled for the on-call center setting. DAGs are represented for the stress factors *Interruptions to family/leisure time* (a), *Anticipating action* (b), *Potential errors* (c), *Hourly caseload* (d), *Post-call recovery* (e), *Sex* (f), *Experience* (g), *Unfair questions* (h), and *Demanding behaviour* (i), respectively

is 2.5 units higher in comparison with doctors experiencing only some interruptions.

2. Doctors who easily recognize a feeling of anticipating future events during their on-call shift experience a stress level that is 2.9 units higher than doctors who only slightly recognize this feeling.
3. Doctors who are only slightly worried about the potential errors they may make during the on-call shift have a stress level that is 2.9 units lower than doctors who are rather worried about potential errors.

Discussion

Main findings and interpretation

We aimed to examine causal effects of specific stress factors of interest based on data from an observational survey. Resting on the assumption that our causal DAGs are correct, we found that stress factors stemming from psychological or emotional sources within the physician, contribute significantly to a feeling of stress or tension. Specifically, the experienced interruptions to leisure time due to an on-call shift, the fear of possibly making medical errors, and anticipatory behaviour in response to future calls or patient visits, emerged as key contributors to elevated stress levels in GPs in training. However, it is necessary to interpret these findings with caution, as a greater variability in participants' responses to other stress factors may limit our ability to measure their significance accurately. Other stress factors, such as patients' questions that are not appropriate for an urgent care setting or demanding behaviour from patients, were highlighted as important factors during group discussion. Additional data collection is required to further investigate these factors.

Comparison of findings with existing literature

The demands of on-call work and its impact on the well-being of physicians have been the focus of other studies as well. Articles encountering GPs or specialist physicians, identified similar stress factors as in our study conducted with GPs in training, such as interference with family or leisure time [6, 10], anticipatory behaviour prior to possible patient contacts [10] and concerns regarding potential errors [20]. For instance, Vincent et al. support our findings, noting that on-call workers in both primary and secondary care mainly ruminated about interruptions to leisure or family time, the risk of missing calls, or were preoccupied with anticipating future calls. Especially female on-call workers appeared to exhibit more anticipatory behaviour, such as thinking about the likelihood of being called and reporting thoughts about pre-planning their visits [10]. Nevertheless, our study did not find any significant effect of gender on the above mentioned stress factors.

Similar to our findings, Zahrai et al. revealed that concerns about medical errors or patient death were the main work-related issues for residents during night shifts, highlighting the need for accessible supervisory support [20]. These results emphasize the need to include physicians' voice in organizational decisions in order to arrange on-call duties according to their specific needs. Finally, Heponiemi et al. explored demographic features among physicians to assess whether there are groups who are more vulnerable to on-call work. Female gender, age and specialization status appeared to be important demographic factors [7]. Since we could not yet demonstrate a significant effect of any demographic factor on occupational stress on-call in our analysis, further research is still needed.

Furthermore, the existing literature identified additional stress factors which, however, were not found to be significant in our study. An increasing hourly caseload [5, 9, 21], an increasing number of shifts per month [9, 16, 22, 23], or number of hours per shift [7, 8, 11, 24] has been associated with mental fatigue, burnout or occupational stress. Also time pressure [7] and little educational value [11] were stated as stress factors, but were not explored in our survey. On the other hand, organizational justice [7], the presence of a recovery period post-call [11, 12], social support [5], a healthy team climate [7, 11] and a supporting supervisor available [7, 13] proved to be effective job resources balancing the experienced on-call demands and stress. Notably, while post-call recovery and hourly caseload were identified as relevant, they did not demonstrate a statistically significant effect on stress in our study. These differences in study results can be attributed to variations in study design (e.g. weekday versus weekend shift) or to differences in on-call systems across countries or within the same on-call system. Also, differences between hospital and primary care settings contribute to these discrepancies. This suggests the need to examine differences in stress factors depending on the specific context, or compared to those found in hospital settings.

Strengths and limitations

Strengths and innovativeness of the study

To our knowledge, there are currently no other studies that utilize DAGs as a methodological strategy to infer causal effects from observational data in the context of out-of-hours care. Although DAGs are increasingly popular as an evidence-based strategy to guide study design, they remain relatively uncommon within the broader scope of applied health research. Tennant et al. [18] identified a knowledge gap in developing DAGs and highlighted the need for practical guidance. While mastering the use and development of DAGs demands some training, their scientific merit is well established in

the available evidence. DAGs encouraged critical reflection on our research questions and offered a transparent structure to integrate insights from literature, group discussion and domain experts. As noted by Fleischer et al. [25], the strength of DAGs lies in making simplifications of reality explicit and open to scrutiny. Finally, DAGs provide a formal framework to control for confounding in causal analysis. In conclusion, the use of DAGs has contributed to the 'Open Science Movement', which emphasizes research quality and innovation characterized by transparency, rigor, reproducibility and replicability [26].

In addition to the methodological innovation, the study's novelty also lies in its population. Until now, most research has focused primarily on general practitioners or specialized hospital physicians. Since there has been growing attention to the existing stress factors of young doctors, we have specifically tailored our research question to GPs in training who frequently undertake on-call shifts.

Limitations

The limitations of our study include the following. First, the literature review does not meet all criteria required for a systematic literature review, as it was restricted to a limited scientific database. We used primarily PubMed with supplementary use of Google Scholar, in contrast to the current recommendation to involve multiple relevant databases, as this increases the reliability of the assessment. Since it was not our primary goal to conduct a systematic review, we found this to be a pragmatic strategy. This allowed to obtain a comprehensive overview to approximate saturation on the most often published risk factors for occupational stress during out-of-hours care. Readers should keep in mind that the inclusion of a certain variable in our causal diagram may be based on limited evidence. Second, the sample size used in the survey was rather small, considering we worked with categorical variables which typically necessitate larger datasets compared to continuous variables. Third, our sample consisted of GPs in training who voluntarily chose to participate, potentially introducing selection bias.

Implications for policy and clinical practice

The results of our study can point towards possible improvements of clinical practice and education of GPs in training, prioritizing their mental health. Given the finding that internal stress factors have proven to be significant predictors of the occupational stress experienced during the on-call shift, there is a need to focus on coping mechanisms and resilience training in physicians' education. This finding underlines the importance

of maintaining a healthy work-life balance to be able to commence an on-call shift in balanced, well-rested mental condition, an aspect that should already receive considerable attention during medical school. To inform decisions to be made at the regulatory level, it would be wise to conduct a larger follow-up study focusing on those stress factors, stemming from psychological or emotional sources within the physician.

Multiple strategies to tackle such internal stress factors may be devised. Firstly, the presence of an external confidential contact person with whom physicians can share threatening encounters or mental health problems, might be valuable to explore in future research. Moreover, the merit of a triage system was repeatedly mentioned during the group discussion. A triage system can function as a filter to avoid unnecessary patient encounters and inappropriate questions, thereby reducing the caseload during on-call shifts, and thus associated anticipatory actions. As a result, physicians would be less fatigued after their shift, experiencing less interruptions to leisure time.

Literature suggests that the competitive nature of medical school, imposing a perfectionist attitude and a performance pressure, creates a lot of stress during the clinical years [27]. As such, we might cautiously hypothesize that such an educational culture might contribute to the fear of making potential errors or to anticipatory behavior. Further research is needed to investigate this hypothesis. Additionally, since GPs in training perform a substantial amount of out-of-hours care, it would be helpful if they had a voice in policy making and organization of the on-call system. This way, important stressors and preventive strategies might be revealed.

During the group discussion, participants agreed on the need to educate patients to enhance their health literacy and knowledge about the appropriate use of the out-of-hours service. Providing patients with guidelines on when to visit the emergency room or a GP cooperative on-call center, will reduce the number of patient visits and unnecessary concerns. If possible, such health skills or feedback on the appropriateness of the visit could be provided to patients after their visit. Finally, as physicians, we should make clear guidelines across different GP cooperative on-call centers, regarding which advices or practices can be performed in an on-call setting to achieve uniformity and clarity to patients.

Conclusion

This study presents an investigation into the causal effects of key stress factors experienced by GPs in training during an on-call shift. We found that internal psychological

and emotional factors were the main contributors to occupational stress in our study, emphasizing the need to raise awareness about the challenges faced by GPs in training, and to promote effective preventive measures in healthcare.

Appendix 1: Search strategy

The following terms were used in Mesh and free-text searches, targeting two main topics:

1. “On-call modalities”[Mesh] OR “Night call”[Mesh] OR “Weekend call”[Mesh] OR “Out-of hours primary care”[Mesh]
2. “Stress factors”[Mesh] OR “Stress”[Mesh] Inclusion criteria: The inclusion criteria comprised studies evaluating the impact of stress factors, whether shift-related or personal to healthcare professionals, on the outcome measure of occupational stress during on-call work. Given the scarcity of results, we included studies conducted with both primary care and specialist physicians (emergency medicine, internal medicine and surgical or interventional disciplines). Studies with nurses or other paramedical professions were excluded. We considered various on-call modalities such as a night float system, a traditional 24 h call or an evening call. The outcome measure was restricted to occupational stress experienced during or in the first days following the on-call shift.

Selection and Data Extraction: Study selection was performed using a multi-stage approach. First, we checked titles to assess whether the articles were valuable to the research question. Subsequently, the abstracts of the remaining studies were examined and any articles that were not appropriate to answer the research question or failed to meet the inclusion criteria, were eliminated at this stage. Finally, all reports of studies identified as potentially eligible were assessed in full text. Additionally, we reviewed the reference lists of identified publications for other relevant studies that may have been missed previously.

Appendix 2: Buildup of the group discussion

Participants were briefed on the procedure and purpose of the group discussion, as well as on the anonymity and confidentiality of data processing. The discussion was conducted in Dutch and consisted of a welcoming word, an introduction of the study protocol, a warm-up exercise and four open-ended questions scheduled for 1 hour in total. The detailed buildup of the group discussion

in depicted in Table 4 in Appendix 2. Participants were encouraged to share their clinical experiences. The first author led the discussion, supported by an external moderator who paid attention to nonverbal cues and dynamics to ensure that no opinion was overlooked. The discussion was voice recorded and transcribed verbatim afterwards. Eventually, the transcript was coded via NVIVO³, with similar stress factors assigned the same code. This enabled the identification of all mentioned stress factors related to on-call work in the group discussion.

Table 4 Overview of the used structure in the group discussion

Phase	Duration	Discussion
Phase 1: Introduction and informed consent	5 min	
Phase 2: Opening question/icebreaker	5 min	‘It is Friday evening. You look at your schedule and notice that you’re on call on Sunday from 8 am to 7 pm. What is the first thought that crosses your mind?’
Phase 3 : Key questions	10 min	‘What are the most tiring or demanding aspects of being on call for you?’ ‘What are your primary concerns during on-call shifts?’
	10 min	‘Literature review has shown that following factors induce stress during on-call duties. What are the most significant ones for you, and why?’
	10 min	‘There are various types of on-call duty shifts across Europe. Which type of on-call duty would feel the most challenging for you and why?’ ‘Which type of duty would feel the safest for you and why?’
	10 min	‘If you were allowed to change one thing about the organization of the on-call schedule, what would help to reduce the stress?’
Phase 4: Closing question	5 min	‘Have we overlooked anything during the discussion?’ ‘Does anyone want to add anything to this?’
Phase 5: Closing remarks	5 min	

³<https://lumivero.com/products/nvivo/>

Appendix 3: Survey assessing the on-call stress factors

Table 5 Questionnaire assessing on-call stress factors with corresponding questions used in the survey

Assessed on-call stress factors	Corresponding assessment questions used in the survey
Hourly caseload	How busy did you experience the on-call shift?
Medical support	How well did you feel medically supported during the on-call shift (by supervisor, colleagues,...)?
Post-call recovery	If the day after your on-call shift is a working day... Do you have to work the day after your on-call shift, or do you have recovery time?
Experience	In which of the three years of specialized GP training are you currently?
Sex	What is your gender?
Anticipating action	To what extent did you recognize any anticipatory actions during the on-call shift? Examples include: <i>'I constantly keep an eye on my phone,' 'Between calls, I usually cannot relax well but remain alert,' 'I anticipate upcoming patient consults and think about the complaints,' 'During on-call shifts, I cannot sleep well, even if I am not called.'</i>
Interruptions to family time	To what extent did you engage in fewer activities with family or friends as a result of being on-call?
Interruptions to leisure time	To what extent did you experience interruptions in leisure time due to this on-call shift?
Potential errors	To what extent did you worry about making medical errors or potentially missing serious diagnoses during the on-call shift?
Unknown medical records	What was your level of access to the patient's medical records (e.g. medical history, medication list, and allergies...)?
Unfair questions	In your opinion, were patients' chief complaints justified for an urgent on-call setting?
A driver	If a driver was present, how useful was the driver?
A secretary's office	If a secretary's office was present, how useful was the secretary?
Unsafe situations	Did you feel unsafe during this on-call shift due to threatening, aggressive, or demanding patient interactions?
Practical barriers	Did you experience any practical/technical issues during this on-call shift? <i>For example, in processing technical examinations (nasal swabs, blood samples, urine samples...) or completing the payment procedure.</i>
Demanding behaviour	In your opinion, how demanding were the patients during the on-call shift? <i>Examples include patients with excessively high expectations or patients who were critical or suspicious towards your recommendations as a physician.</i>
Medical triage	If a medical triage was present, how useful was the medical triage?
Tolerance level	How did you feel when you left home for the on-call shift?
Hospital resources	How well did you feel supported by the secondary care during the on-call shift? This entails not only the availability of technical resources such as a lab/rapid tests or imaging, but also the intercollegiate collaboration with emergency physicians or specialists.
Geography	In which region did you perform this on-call shift?
SES	In your opinion, what seemed to be the average SES of the patients at the on-call center?
Work environment	Where does the on-call shift take place?

The questionnaire used in the survey was administered in Dutch, the respondents' native language, and has been translated here from Dutch to English

Appendix 4: TBI-CareQOL Caregiver Strain

Table 6 TBI-CareQOL Caregiver Strain – Short Form 6a [16]

	Always	Never	Rarely	Sometimes	Usually
I feel that there are not enough hours in the day to get everything done.	1	2	3	4	5
I feel stressed when it comes to providing care for the person with the injury.	1	2	3	4	5
I feel that stress makes it difficult to start anything new.	1	2	3	4	5
I feel drained by my responsibilities as a caregiver.	1	2	3	4	5
I cannot handle any more responsibility when it comes to providing care for the person with the injury.	1	2	3	4	5
I feel burdened with the responsibility of caring for the person with the injury.	1	2	3	4	5

Appendix 5: Summary of literature review

Table 7 Summary of literature review

Nr.	Factor	Author (year)	Main findings	Study design	Statistics
1.	Hourly caseload	Çalışkan Tür, F. et al. (2016) [5]	Number of patients taken care of by physicians working night shifts was found as a significant job stressor.	Cross-sectional questionnaire study	$p < 0.05$
		Davenport, M. S. et al. (2010) [21]	Increasing cases/hour were significantly and independently associated with increased resident-faculty discrepancy rates.	Prospective cohort study	$p < 0.001$ Odds Ratio = 1.6 [1.5 – 1.7]
		Zhou, Y. et al. (2017) [22]	The number of calls per night shift was significantly associated with burnout.	Cross-sectional questionnaire study	$p = 0.035$
2.	Post-call recovery	Scott, O. et al. (2019) [11]	Data suggest that the extent of the off-call period is an important determinant of burnout and is instrumental in maintaining both mental and physical well-being.	Qualitative research with focus group discussion	NA
		Rose, M. et al. (2008) [12]	Residents were significantly sleepier after an on-call shift. Fatigue continued to recover the second and third day after the on-call shift.	Prospective cohort study	$p < 0.05$
3.	Sex	Çalışkan Tür, F. et al. (2016) [5]	Job stress was higher in female physicians.	Cross-sectional questionnaire study	$p < 0.001$
		Saksvik, I. et al. (2011) [28]	In a recent review, it was concluded that male shift workers tend to sleep better and experience less fatigue and sleepiness when working shifts than women.	Systematic review	$p < 0.05$
		Heponiemi, T. et al. (2015) [7]	Women had more time pressure and stress related to team work compared with men. Women had lower levels of job control and organizational justice than men.	Cross-sectional questionnaire study	$p < 0.001$
		Zhou, Y. et al. (2017) [22]	In a multi-variable analysis, female gender was found to be an independent factor associated with burnout in the pulmonary physicians in this study.	Cross-sectional questionnaire study	$p = 0.018$
		Vincent, G. E. et al. (2021) [10]	The odds of workers pre-planning and thinking about interruptions to leisure/family time were 2 times higher for females than males.	Cross-sectional questionnaire study	$p < 0.05$
4.	Experience	Çalışkan Tür, F. et al. (2016) [5]	Statistically significant weak negative correlation was found between age and job strain.	Cross-sectional questionnaire study	$p < 0.05$
		Vincent, G. E. et al. (2021) [10]	The odds of workers pre-planning were 88% lower for those aged 35-44 years than those aged 18-24, 86% lower for those aged 45-54 years than those aged 18-24 years, and 84% lower for those >55 years than those aged 18-24 years.	Cross-sectional questionnaire study	$p < 0.05$
		Heponiemi, T. et al. (2015) [7]	It has been found that younger workers have better shift work tolerance regarding subjective sleepiness, performance tests, recovery after work, and sleep time. Older respondents had higher levels of stress related to patient information systems and teamwork than younger respondents.	Cross-sectional questionnaire study	$p = 0.006$
5.	Anticipating action	Vincent, G. E. et al. (2021) [10]	On-call workers rated pre-planning in case of a call and fear of missing a call as main reasons for rumination.	Cross-sectional questionnaire study	$p < 0.05$
6.	Interruptions to family/leisure time	Vincent, G. E. et al. (2021) [10]	On-call workers rated interruptions to family/leisure time as one of the main reasons for rumination.	Cross-sectional questionnaire study	$p < 0.05$
		Heponiemi, T. et al. (2014) [6]	Work interference with family was significantly associated with all the outcomes (distress, job satisfaction and work ability).	Prospective cohort study	$p < 0.05$
7.	Potential errors	Zahrai, A. et al. (2011) [20]	With regard to work-related stressors, "concern regarding potential errors" was the only significant work-related problem at 6 months for residents on the night float program.	Prospective cohort study	$p = 0.022$

We retained seven stress factors as most prominent after literature review, outlined with corresponding article references, study findings, design and statistics

Abbreviations

ANOVA	Analysis of Variance
CI	Confidence Interval
DAG	Directed Acyclic Graph
GP	General Practitioner
SES	Socio-Economical State

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12875-025-03140-1>.

Supplementary Material 1.

Acknowledgements

We thank Dr. Franky D'Argent and Anke Boone for their valuable assistance in organizing and guiding our group discussion. Furthermore, we express our gratitude to Dr. Louise Poppe, who provided valuable feedback on the development and application of DAGs, and Prof. Dr. Stefan Heytens, who reviewed the DAGs and provided expert insight. Finally, we are thankful to Heidelinde Dehaene and Alexander Decruyenaere for sharing their insights and expertise regarding the methodology and statistical analysis.

Authors' contributions

GD, PR, SVD, LQ, PD, and TD contributed to the study concept and were responsible for designing the study protocol. GD conducted the literature review, organized the group discussions, and administered the questionnaire survey. PR and SVD performed the data analysis and causal analysis. GD, PR, SVD, LQ, PD, and TD were involved in data interpretation and have read and approved the final version of the manuscript.

Funding

Paloma Rabaey was funded by the Research Foundation Flanders (FWO-Vlaanderen) with grant number 1170124N. Additionally, this research received funding from the Flemish government under the "Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen" program.

Data availability

The authors confirm that the data supporting the findings of this study are available within the article and its supplementary materials.

Declarations

Ethics approval and consent to participate

The study protocol was approved by the Ethics Committee of U/UZ Gent on the 16th December of 2022 (registration number: B6702022000712) in accordance with the Belgian federal law of 7 May 2004 on experiments on the human person. All individual participants involved in the observational study and group discussion provided informed consent to participate.

Consent for publication

This study complies with ethical standards for the use of individual data. All participants in the group discussions and survey provided informed consent, ensuring their agreement to the use of their responses in this research. Data from the questionnaires were pseudonymized to protect the privacy of the respondents.

Competing interests

The authors declare no competing interests.

Received: 11 December 2024 / Accepted: 2 December 2025

Published online: 23 December 2025

References

1. Cox J. Understanding doctors' performance. Radcliffe Publishing; 2006.
2. Tysen R, Vaglum P, Grønvold NT, Ekeberg Ø. Suicidal ideation among medical students and young physicians: a nationwide and prospective study of prevalence and predictors. *J Affect Disord*. 2001;64(1):69–79. [https://doi.org/10.1016/S0165-0327\(00\)00205-6](https://doi.org/10.1016/S0165-0327(00)00205-6).
3. Mata DA, Ramos MA, Bansal N, Khan R, Guille C, Di Angelantonio E, et al. Prevalence of Depression and Depressive Symptoms Among Resident Physicians: A Systematic Review and Meta-analysis. *JAMA*. 2015;314(22):2373–83. <https://doi.org/10.1001/jama.2015.15845>.
4. Dyrbye L, Shanafelt T. A narrative review on burnout experienced by medical students and residents. *Med Educ*. 2016;50(1):132–49. <https://doi.org/10.1111/medu.12927>.
5. Tur FC, Toker I, Sasmaz CT, Hacar S, Ture B. Occupational stress experienced by residents and faculty physicians on night shifts. *Scand J Trauma Resuscitation Emerg Med*. 2016;24:34. <https://doi.org/10.1186/s13049-016-0225-4>.
6. Heponiemi T, Puttonen S, Elovainio M. On-call work and physicians' well-being: testing the potential mediators. *Occup Med*. 2014;64(5):352–7. <https://doi.org/10.1093/occmed/kqu036>.
7. Heponiemi T, Aalto AM, Pekkarinen L, Siuvert E, Elovainio M. Are there high-risk groups among physicians that are more vulnerable to on-call work? *Am J Emerg Med*. 2015;33(5):614–9. <https://doi.org/10.1016/j.ajem.2015.01.034>.
8. Sephien A, Reljic T, Jordan J, Prida X, Kumar A. Resident duty hours and resident and patient outcomes: Systematic review and meta-analysis. *Med Educ*. 2023;57(3):221–32. <https://doi.org/10.1111/medu.14943>.
9. Zhou Y, Gu W, Tao R, Chen C. The impact of gender and working hours on pulmonary physician burnout. *Ann Transl Med*. 2020;8(18):1166. <https://doi.org/10.21037/atm-20-5827>.
10. Vincent GE, Kovac K, Signal TL, Reynolds AC, Paterson J, Sprajcer M, et al. What Factors Influence the Sleep of On-call Workers? *Behav Sleep Med*. 2021;19(2):255–272. Epub 2020 Feb 27. <https://doi.org/10.1080/15402002.2020.1733575>.
11. Scott O, Novak C, Forbes K. Medical Student Perceptions of On-Call Modalities: A Focus Group Study. *Teach Learn Med*. 2019;31(1):34–43. <https://doi.org/10.1080/10401334.2018.1480957>.
12. Rose M, Manser T, Ware JC. Effects of call on sleep and mood in internal medicine residents. *Behav Sleep Med*. 2008;6(2):75–88. <https://doi.org/10.1080/15402000801952914>.
13. Jasti H, Hanusa BH, Switzer GE, Granieri R, Elnicki M. Residents' perceptions of a night float system. *BMC Med Educ*. 2009;9:52.
14. Mortelmans D. Handboek kwalitatieve onderzoeksmethoden. Acco Leuven; Blijde Inkomststraat, 3000, Leuven; 2009.
15. Steenhuyzen S, Stevens V. Een kei in evalueren: Werkboek Evalueren in de gezondheidsbevordering. Antwerpen-Appeldoorn: Garant; 2016.
16. Alenezi K, Alyami H, Alrehaili O, Arruhaily A, Alenazi K, Al-Dubai AR. Prevalence and Associated Factors of Burnout Among Saudi Resident Doctors: A Multicenter Cross-Sectional Study. *Alpha Psychiatry*. 2022;23(4):173–83. <https://doi.org/10.5152/alphapsychiatry.2022.21361>.
17. Steeman L, Uijen M, Plat E, Huibers L, Smits M, Giesen P. Out-of-hours primary care in 26 European countries: an overview of organizational models. *Fam Pract*. 2020;37(6):744–50. <https://doi.org/10.1093/fampra/cmaa064>.
18. Tennant PW, Murray EJ, Arnold KF, Berrie L, Fox MP, Gadd SC, et al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. *Int J Epidemiol*. 2021;50(2):620–32. <https://doi.org/10.1093/ije/dyaa213>.
19. Poppe L, Steen J, Loh W, Crombez G, De Block F, Jacobs N, et al. How to develop causal directed acyclic graphs for observational health research: a scoping review. *Health Psychol Rev*. 2024;1–21. Epub ahead of print. <https://doi.org/10.1080/17437199.2024.2402809>.
20. Zahrai A, Chahal J, Stojimirovic D, Schemitsch EH, Yee A, Kraemer W. Quality of life and educational benefit among orthopedic surgery residents: a prospective, multicentre comparison of the night float and the standard call systems. *Can J Surg*. 2011;54(1):25–32. Erratum in: *Can J Surg*. 2011;54(2):88.
21. Davenport MS, Ellis JH, Khalatbari SH, Myles JD, Klein KA. Effect of work hours, caseload, shift type, and experience on resident call performance. *Acad Radiol*. 2010;17(7):921–7. <https://doi.org/10.1016/j.acra.2010.03.006>.

22. Zhou X, Pu J, Zhong X, Zhu D, Yin D, Yang L, et al. Burnout, psychological morbidity, job stress, and job satisfaction in Chinese neurologists. *Neurology*. 2017;88(18):1727–35.
23. Yasin R, Muntham D, Chirakalwasan N. Uncovering the sleep disorders among young doctors. *Sleep Breath*. 2016;20(4):1137–44. <https://doi.org/10.1007/s11325-016-1380-6>.
24. Mautone S. Toward a new paradigm in graduate medical education in the United States: elimination of the 24-hour call. *J Grad Med Educ*. 2009;1(2):188–94. <https://doi.org/10.4300/JGME-D-09-00061.1>.
25. Fleischer NL, Diez Roux AV. Using directed acyclic graphs to guide analyses of neighbourhood health effects: an introduction. *J Epidemiol Community Health*. 2008;62(9):842–6. <https://doi.org/10.1136/jech.2007.067371>.
26. UNESCO. An introduction to the UNESCO Recommendation on Open Science. UK National Commission for UNESCO, Paris; 2022.
27. Morcos G, Awan OA. Burnout in Medical School: A Medical Student's Perspective. *Acad Radiol*. 2023;30(6):1223–5. <https://doi.org/10.1016/j.jacr.2022.11.023>.
28. Saksvik IB, Bjorvatn B, Hetland H, Sandal GM, Pallesen S. Individual differences in tolerance to shift work-a systematic review. *Sleep Med Rev*. 2011;15(4):221–35.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.