BMJ Open

# Factors associated with under- and overtriage in telephone triage in Danish out-of-hours primary care – a natural quasi-experimental cross-sectional study of randomly selected and high-risk calls

| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2022-064999 |
| Article Type: | Original research |
| Date Submitted by the Author: | 26-May-2022 |
| Complete List of Authors: | Graversen, Dennis; Aarhus University Research Unit General Practice; Aarhus University, Department of Public Health<br>Pedersen, Anette; Aarhus University Research Unit General Practice; Aarhus University, Department of clinical medicine<br>Christensen, Morten; Aarhus University Research Unit General Practice; Aarhus University, Department of Public Health<br>Folke, Fredrik; Copenhagen University Hospital, Emergency Medical Services Copenhagen; Copenhagen University Hospital, Department of Cardiology - Herlev and Gentofte<br>Huibers, L; Aarhus University Research Unit General Practice, Research Unit for General Practice |
| Keywords: | Clinical audit < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health & safety < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, PRIMARY CARE |
| | |

*I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.*

*The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.*

*Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.*

# Factors associated with under- and overtriage in telephone triage in Danish out-of-hours primary care – a natural quasi-experimental cross-sectional study of randomly selected and high-risk calls

**Graversen DS[1,2], Pedersen AF[1,3], Christensen MB[1,2], Folke F[4,5,6], Huibers L[1]**

[1]Research Unit for General Practice, Aarhus, Denmark, [2]Department of Public Health, Aarhus University, Aarhus, Denmark, [3]Department of Clinical Medicine, Aarhus University, Aarhus, Denmark, [4]Copenhagen University Hospital - Emergency Medical Services Copenhagen, Ballerup, Denmark, [5]Department of Cardiology, Copenhagen University Hospital - Herlev and Gentofte, Denmark, [6]Department of Clinical Medicine, University of Copenhagen, Denmark

*Correspondence to*

DS Graversen (ORCID ID: 0000-0002-7784-9403), Research Unit for General Practice, Aarhus, Bartholins Allé 2, 8000 Aarhus C, Denmark & Department of Public Health, Aarhus University, Bartholins Allé 2, 8000 Aarhus C, Denmark

E-mail: d.graversen@ph.au.dk, telephone: +45 28 26 05 52

Keywords: Triage, Telephone, After-Hours Care, Primary Health Care, General Practitioners, Nurses, Safety, Efficiency, Quality of Health Care

Abstract: words 298

Word count: 2676

**Abbreviations:**

AQTT: Assessment of Quality in Telephone Triage (a validated tool for assessment of triage performance)

CDSS: Computerised decision support system

GP: General practitioners

GPC: General practitioners cooperative

MH-1813: Medical helpline 1813

OOH: Out-of-hours

OOH-PC: Out-of-hours primary care

**ABSTRACT:**

**Objectives**: We aim to explore under- and overtriage in high-risk patient population and explore patient- and call characteristics associated with under- and overtriage in both randomly selected and in high-risk telephone calls to out-of-hours primary care (OOH-PC).

**Design:** Natural quasi-experimental cross-sectional study.

**Setting:** Two Danish out-of-hours (OOH) services using different telephone triage models: a general practitioner cooperative (GPC) with GP-led triage and the medical helpline 1813 with nurse-led triage.

**Participants:** We included audio-recorded telephone triage calls from 2016: 806 random calls and 405 high-risk calls (defined as patients ≥ 30 years calling with abdominal pain).

**Main outcome measures:** Twenty-four experienced physicians used a validated assessment tool to assess the accuracy of triage (AQTT). We calculated the relative risk (RR) for *clinically relevant* under- and overtriage for a range of patient- and call characteristics.

**Results:** We included 806 randomly selected calls (44 *clinically relevant* undertriaged and 54 *clinically relevant* overtriaged) and 405 high-risk calls (32 undertriaged and 24 overtriaged). In high-risk calls, nurse-led triage was associated with significantly less undertriage (RR: 0.47, 95%CI: 0.23-0.97) and more overtriage (RR: 3.93, 95%CI: 1.50-10.33) compared to GP-led triage. In high-risk calls, the risk of undertriage was significantly higher for calls during nighttime (RR: 2.1, 95%CI: 1.05-4.07). Undertriage tended to be more likely for calls concerning patients ≥ 60 years compared to 30 to 59 years (11.3% vs. 6.3%) in high-risk calls. However, this result was not significant. The risk of *clinically relevant* under- and overtriage was similar in randomly selected and high-risk calls.

**Conclusion:** Nurse-led triage was associated with less undertriage and more overtriage compared to GP-led triage in high-risk calls. This study may suggest that to minimize undertriage, the triage professionals should pay extra attention when a call occurs during nighttime or concerns elderly. However, this needs confirmation in future studies.

## Article Summary

### Strengths and limitations of this study

- A strength of our study is the natural quasi-experimental design using real-life calls as opposed to the constructed setup.
- To our knowledge, this is the first study to explore factors associated with assessed undertriage and overtriage in real audio-recorded calls.
- A limitation is the use of only one assessor per call, but acceptable interrater agreement was previously found.

## INTRODUCTION

In out-of-hours primary care (OOH-PC), telephone triage plays a pivotal role in managing patient flows and workload[1–3]. Telephone triage aims to ensure a safe and efficient delivery of healthcare, avoiding undertriage and minimizing overtriage[4]. However, accurate telephone triage is a challenge due to the lack of visual cues of the patient, challenging communication, and time pressure[5–8].

Safety and efficiency of telephone triage in OOH-PC have been explored in a range of studies that used varying outcome measures [4,9–13]. Unsafe telephone triage in out-of-hours (OOH) care has been associated with calls concerning abdominal pain[14–20], chest pain[16,17] and shortness of breath[16,18], calls for patients with increasing age[14,17,21], and calls during nighttime[17,22]. A few studies have explored factors associated with overtriage as a measure of inefficient telephone triage. Nurse-led triage [1,23] and triage aided by computerized decision support system (CDSS)[23–26] have been associated with inefficient telephone triage. Moreover, several studies suggested that calls concerning children are difficult to triage[22,27–31].

Previously, we found that nurse-led triage was associated with significantly less undertriage and significantly more overtriage compared to GP-led triage in a sample of random calls[23] . However, it remains unclear whether this association also exists for high-risk calls. Therefore, we aim to explore the risk of under- and overtriage in a selected population of high-risk calls. In addition, we aim to explore patient- and call characteristics associated with undertriage and overtriage in telephone triage in both randomly selected calls and in a selection of high-risk calls in OOH-PC.

## METHODS

*Design and setting*

This paper presents secondary analyses of a natural quasi-experimental study in two regional OOH-PC services in Denmark[23,32]. Telephone triage in OOH-PC is provided by the general practitioner cooperative

(GPC) in the Central Denmark Region, and by the medical helpline 1813 (MH-1813) the Capital Region of Denmark. Both services are open outside office hours (i.e., on weekdays from 4 pm to 8 am, weekends, and national holidays), offering telephone consultations, clinic consultations, and home visits. The GPC uses GP-led telephone triage, whereas the MH-1813 uses nurse-led telephone triage (see Box 1). At the MH-1813, registered nurses are obliged to use a locally developed CDSS and they can redirect calls to a physician on call. These physicians with various medical speciality also conduct telephone triage. In the present paper, we excluded calls by physicians at MH-1813 from analyses.

**Box 1.** Description of the organisation and telephone triage of the two included models for out-of-hours primary care

| | **GP cooperative (GPC)** | **Medical helpline 1813 (MH-1813)** |
|---|---|---|
| **Region** | **Central Denmark Region** | **Capital Region of Denmark** |
| **Population**[33] | 1.3m citizens | 1.8m citizens |
| **Telephone calls in 2014**[34] | 697,000 | 911,000 |
| **Organiser** | GPs in the region | Regional administration |
| **Organisation and services** | ▪ Telephone triage, home visits, and face-to-face consultations at the GPC<br>▪ GPs are obliged to take part in the service | ▪ Telephone triage and home visits run by MH-1813<br>▪ Face-to-face consultations are located in hospital facilities and managed by EDs |
| **Remuneration of professionals** | Fee for service | Payment by the hour |

| | | |
|---|---|---|
| **Triage professional** | GPs or GP trainees in their final year of speciality; no CDSS available | Nurses who are obliged to use a CDSS and have option to redirect calls to a physician<br>Physicians with different medical specialities (a minority being GPs) |

*CDSS: Computerized decision support system*



*Definition of groups of calls*

We defined two groups of calls: 1) a group of randomly selected calls that was representative for all calls to OOH-PC and 2) a group of potentially high-risk patient calls (referred to as "high-risk"). To define the high-risk patient calls, we conducted a systematic literature search in 2016 to identify factors associated with unsafe or inaccurate telephone triage. Seven identified studies described factors associated with a risk of unsafe telephone triage: infants[22], increasing patient age[14,17,21], calls during nighttime[17,22], abdominal pain[14–18], chest pain[16,17], and breathing difficulties[16,18]. In a consensus meeting, the authors (DSG, AFP, MBC, LH,) defined criteria for high-risk calls as calls concerning patients above 30 years who suffered from abdominal pain. We added the age criterion, as we aimed to include potentially dangerous conditions that present with vague, indistinctly, or greatly differing symptomatology (such as dissecting aorta aneurism, myocardial infarction, cholecystitis, pyelonephritis, acute pancreatitis, and gastrointestinal ulcer).

*Selection of calls*

All calls to the GPC and MH-1813 outside office hours during the inclusion period (GPC: 23 November - 7 December 2016; MH-1813: 23 November - 8 December 2016) were potentially eligible. For calls redirected by a nurse to a physician at MH-1813, only the part conducted by the nurse was eligible and later assessed as described below. Randomly selected calls were selected from all eligible calls. The aim was to include

435 calls by GPs and 435 by nurses, based on a power calculation to detect differences in undertriage[23]. Based on expected exclusion, we randomly selected 525 random calls from GPC and 500 calls from MH-1813.

To identify potential eligible high-risk calls, we received electronic patient journal records from the GPC and MH-1813. We defined a list of wordings and abbreviations associated with abdominal pain, which was used to search the patient record and identify all eligible calls concerning including abdominal pain. DSG assessed these marked patient records to check whether inclusion for abdominal pain was correct. Calls were excluded when the triage professional in the patient record dismissed the presence of abdominal pain (i.e., triagists noted *no abdominal pain*) or when the complaint was clearly outside the thoracic, abdominal, or pelvic region. Thus, we identified 846 (GPC) and 884 (MH-1813) eligible high-risk calls for potential inclusion, of which we randomly selected 252 calls from the GPC and 240 calls from MH-1813. Selection of both groups of calls was done with Stata, matching the overall distribution on day of week (i.e., weekend, not weekend) and time of day (i.e., day, evening, night).

Each selected call had a unique identification number that was used to identify the corresponding audio recording. Three master students of medicine masked the audio recordings using beep tones to cover information revealing triage profession, OOH organisation, and patient identification information. In addition, the students screened all calls for fulfilling exclusion criteria as shown in Appendix 1. DSG reviewed all calls that fulfilled exclusion criteria and all calls for which a student was uncertain. In case of doubt, DSG and AFP reached consensus regarding exclusion.

*Assessment of accuracy of triage*

All included calls have been assessed as described prior papers[23,35] using the tool "Assessment of Quality in Telephone Triage" (AQTT). The AQTT comprises 24 items assessing the health-related quality and the

quality of communication[35]. One item assessed the accuracy of triage on a 7-point scale to differentiate between levels of undertriage and overtriage according to severity. The AQTT showed inter-rater disagreement when using the entire scale for assessment of the accuracy of triage, but revealed satisfactory inter-rater agreement when distinguishing ratings "1" & "2" (defined as *clinically relevant* undertriage) from ratings "3" to "7" and ratings "6" & "7" (defined as *clinically relevant* overtriage) from ratings "1" to "5"[35].

For the assessment panel, we recruited 24 physicians among triage professionals from the GPC and MH-1813 using two inclusion criteria: 1) >1 year experience and 2) active in telephone triage in OOH-PC at time of study. We randomly selected 16 GPs from the 56 interested GPs from the GPC. At MH-1813, we included all eight physicians fulfilling our inclusion criteria from the ten interested physicians. All assessors followed a two-day training course providing knowledge on telephone triage and communication, introducing the AQTT and rating manual, and assessing triage calls individually and in plenary, focusing on achieving consistency. We randomly distributed calls to all assessors, so each member of the assessment panel assessed random and high-risk calls triaged by GPs and by nurses. Information on age and sex of the patient, day of week, and the time of each call was available for the assessor, extracted from the registration systems from the GPC and MH-1813.

*Statistical analyses*

Accuracy of triage decision was categorised into *clinically relevant* undertriage (rated "1" or "2") and *clinically relevant* overtriage (rated "6" or "7"). We conducted separate analyses for both type of calls (i.e., randomly selected calls and high-risk calls). We used descriptive analyses to describe patient- and call characteristics, as well as the risk of *clinically relevant* undertriage and *clinically relevant* overtriage for both type of calls. To explore differences in the risk of *clinically relevant* under- and overtriage between randomly selected and high-risk calls, we excluded patients <30 years old from randomly selected calls aiming for equal age inclusion criterion. Relative risks (RRs) were calculated using binomial regression

analyses. Next, we explored the association of individual patient- and call characteristics with inaccurate telephone triage, by calculating the RR of having *clinically relevant* undertriage (vs. no *clinically relevant* undertriage) and *clinically relevant* overtriage (vs. no *clinically relevant* overtriage), using binomial regression. 95% confidence intervals (CI) were calculated. All analyses were performed in Stata 14.2 (StataCorp. 2015. *Stata Statistical Software: Release 14.2*. College Station, TX: StataCorp LP).

## PUBLIC AND PATIENT INVOLVEMENT STATEMENT

We explored patients' perspective in a focus group interview concerning the development of the AQTT and incorporated received input in the rating manual of the AQTT [35].

## RESULTS

*Description of calls*

Figure 1 shows the flowchart of selection and exclusion of calls. We excluded 47 randomly selected calls and 30 high-risk calls assigned 'not applicable', as assessing accuracy of triage was not possible (e.g., insufficient information was available) or assessing not relevant. We included 806 randomly selected calls and 405 high-risk calls (Table 1). The risk of *clinically relevant* undertriage was 5.5% in randomly selected calls (including all age groups) and 7.9% in high-risk calls. We found similar risks of *clinically relevant* undertriage (RR=1.04, 95%CI: 0.63-1.72) and *clinically relevant* overtriage (RR=1.07, 95%CI: 0.60-1.92) for randomly selected and high-risk calls (when only including patients ≥30 years to have similar age criteria) (data not shown in Table).

*(Figure 1 – flowchart- attached separate)*

**Figure 1.** Flowchart of inclusion and exclusion of randomly selected and high-risk calls from the GPC and MH-1813

*Not applicable:* accuracy of triage could not be assessed due to insufficient information in the call or assessment of accuracy of triage was deemed not relevant.

**Table 1.** Baseline description of patient and call characteristics in randomly selected calls and high-risk calls

| | **Randomly selected calls** | **High-risk calls[1]** |
|---|---|---|
| | **N=806[2]** | **N=405[2]** |
| ***Patient characteristics*** | % (N) | % (N) |
| **Age groups (in years)** | | |
| - <18 | 36.5 (294) | N.a.[1] |
| - 18-29 | 21.0 (169) | N.a.[1] |
| - 30-59 | 27.5 (222) | 67.2 (272) |
| - ≥60 | 15.0 (121) | 32.8 (133) |
| **Sex** | | |
| - Male | 40.1 (323) | 40.3 (163) |
| - Female | 59.9 (483) | 59.8 (242) |
| ***Call characteristics*** | | |
| **Weekend** | | |
| - Not weekend | 46.4 (374) | 51.4 (208) |
| - Weekend | 53.6 (432) | 48.6 (197) |
| **Time of day** | | |
| - Day or evening | 83.6 (674) | 77.5 (314) |
| - Nighttime | 16.4 (132) | 22.5 (91) |
| **Assessed accuracy of triage[3]** | | |

| | | |
|---|---|---|
| *Clinically relevant* undertriage | | |
| 1: Severe undertriage | 1.5 (12) | 3.0 (12) |
| 2: Moderate undertriage | 4.0 (32) | 4.9 (20) |
| *Satisfactory* triage | | |
| 3: Mild undertriage | 8.6 (69) | 11.6 (47) |
| 4: Optimal triage | 69.0 (556) | 66.2 (268) |
| 5: Mild overtriage | 10.3 (83) | 8.4 (34) |
| *Clinically relevant* overtriage | | |
| 6: Moderate overtriage | 4.6 (37) | 4.4 (18) |
| 7: Severe overtriage | 2.1 (17) | 1.5 (6) |
| **Triage professional** | | |
| - GP | 49.5 (399) | 50.9 (206) |
| - Nurse | 50.5 (407) | 49.1 (199) |

[1]*High-risk calls included patients aged ≥30 years calling OOH-PC with abdominal pain, thus patients aged <30 years were not applicable.* [2]*Not applicable: Calls where the assessment of accuracy of triage was not possible or not relevant was excluded (randomly selected n=47, high-risk n=30).* [3]*Accuracy of triage was categorized for further analyses: 1) satisfactory triage, including optimal ("4") or mild under-("3") and overtriage ('5'), 2) clinically relevant undertriage ('"1" or "2"), and 3) clinically relevant overtriage ("6" or "7").*

*Randomly selected calls*

In randomly selected calls, we found no significant association between patient age or sex, whether call was conducted during weekend, or what time of day the call was conducted and the risk of clinically relevant under- or overtriage (Table 2).

**Table 2.** Patient- and call characteristics associated with clinically relevant under- and overtriage in randomly selected calls[1]

| | Undertriage[2] | | Overtriage[3] | |
|---|---|---|---|---|
| | N=44 | | N=54 | |
| | % (N) | RR[2] (95%CI) | % (N) | RR[2] (95%CI) |
| ***Patient characteristics*** | | | | |
| **Age groups (in years)** | | | | |
| - <18 (n=294) | 3.7 (11) | 0.57 (0.23-1.37) | 7.5 (22) | 1.51 (0.63-3.63) |
| - 18-29 (n=169) | 4.1 (7) | 0.62 (0.23-1.68) | 7.7 (13) | 1.55 (0.61-3.97) |
| - 30-59 (n=222) | 8.1 (18) | 1.23 (0.55-2.74) | 5.9 (13) | 1.18 (0.46-3.03) |
| - ≥ 60 (n=121) | 6.6 (8) | 1 | 5.0 (6) | 1 |
| **Sex** | | | | |
| - Male (n=323) | 5.3 (17) | 1 | 4.6 (15) | 1 |
| - Female (n=483) | 5.6 (27) | 1.06 (0.59-1.92) | 8.1 (39) | 1.74 (0.97-3.10) |
| ***Call characteristics*** | | | | |
| **Weekend** | | | | |

| | | | | |
|---|---|---|---|---|
| - Not weekend (n=374) | 5.4 (20) | 1 | 6.7 (25) | 1 |
| - Weekend (n=432) | 5.6 (24) | 1.04 (0.58-1.85) | 6.7 (29) | 1.00 (0.60-1.68) |
| **Time of day** | | | | |
| - Day or evening (n=674) | 5.3 (36) | 1 | 7.0 (47) | 1 |
| - Nighttime (n=132) | 6.1 (8) | 1.13 (0.54-2.39) | 5.3 (7) | 0.76 (0.35-1.65) |

*[1]We excluded 47 calls of the selected 853 calls prior to analyses, as assessing accuracy of triage was not possible or not relevant.*
*[2]Undertriage is triage decision assessed as clinically relevant undertriage (rated "1" or "2"): [3]overtriage is clinically relevant overtriage (rated "6" or "7") on a Likert scale from 1 to 7;*
**Significant difference in RR, using binomial regression analyses; p<0.05*

*High-risk calls*

Nurse-led triage was associated with significantly less *clinically relevant* undertriage than GP-led triage (RR=0.47 95%CI: 0.23-0.97) (Table 3). Nurse-led triage had significantly more *clinically relevant* overtriage (RR=3.93, 95%CI: 1.50-10.53) compared to GP-led triage. For high-risk calls, calls conducted during nighttime were associated with a significantly higher risk of being *clinically relevant* undertriaged (13.2%) compared to calls conducted during day or evening (6.4%) (RR=2.1, 95%CI: 1.05-4.07). We found no significant associations in any of the other patient- and call characteristics. However, a close to significant trend was seen for patient age, as the risk of *clinically relevant* undertriage was lower in patients aged 30 to 59 years (6.3%) in comparison with elderly patients ≥60 years (11.3%) (RR=0.55, 95%CI: 0.29-1.07).

**Table 3.** Patient- and call characteristics associated with clinically relevant under- and overtriage in high-risk calls

| | **Undertriage[1]** | | **Overtriage[2]** | |
|---|---|---|---|---|
| | **N=32** | | **N=24** | |
| | % (N) | RR (95%CI) | % (N) | RR (95%CI) |
| ***Patient characteristics*** | | | | |
| **Age groups (in years)** | | | | |
| - 30-59 (n=272) | 6.3 (17) | 0.55 (0.29-1.07) | 7.4 (20) | 2.44 (0.85-7.01) |
| - >60 (n=133) | 11.3 (15) | 1 | 3.0 (4) | 1 |
| **Sex** | | | | |
| - Male (n=163) | 8.6 (14) | 1 | 6.8 (11) | 1 |
| - Female (n=242) | 7.4 (18) | 0.87 (0.44-1.69) | 5.4 (13) | 0.80 (0.40-1.73) |
| ***Call characteristics*:** | | | | |
| **Weekend** | | | | |
| - Not weekend (n=208) | 6.7 (14) | 1 | 6.3 (13) | 1 |
| - Weekend (n=197) | 9.1 (18) | 1.36 (0.69-2.65) | 5.6 (11) | 0.89 (0.41-1.95) |
| **Time of day** | | | | |
| - Day or evening (n=314) | 6.4 (20) | 1 | 6.1 (19) | 1 |
| - Nighttime (n=91) | 13.2 (12) | 2.1 (1.05-4.07)* | 5.5 (5) | 0.91 (0.35-2.36) |

| | | | | |
|---|---|---|---|---|
| **Triage professional** | | | | |
| - GP (n=206) | 10.7 (22) | 1 | 2.4 (5) | 1 |
| - Nurse (n=199) | 5.0 (10) | 0.47 (0.23-0.97)* | 10.0 (19) | 3.93 (1.50-10.33)* |

*We excluded 30 calls as assessing accuracy of triage was not possible or not relevant.*
*[1]Undertriage is triage decision assessed as clinically relevant undertriage (rated "1" or "2"): [2]overtriage is clinically relevant overtriage (rated "6" or "7");*
**Significant difference in RR, using binomial regression analyses; p<0.05*

## DISCUSSION

### *Principal findings*

In high-risk calls, nurse-led triage was associated with significantly lower risk of *clinically relevant* undertriage and significant higher risk of *clinically relevant* overtriage compared to GP-led triage. The risk of *clinically relevant* undertriage and *clinically relevant* overtriage was similar for randomly selected and high-risk calls. In high-risk calls, the risk of *clinically relevant* undertriage was significantly higher if the call was conducted during nighttime compared to day and evening.

### *Strengths and limitations of the study*

A major strength was the quasi-experimental design using real-life calls as opposed to the constructed setup used in previous studies [18–20, 31, 32]. An additional strength was the meticulous assessment process using the validated AQTT tool combined with a comprehensive rating manual [35].

Our study had some limitations. Due to the thorough assessment process, each call was assessed by one assessor. Consequently, bias due to misclassification cannot be rejected as some subjectivity may remain. However, we took several precautions to ensure consistency of assessments; the assessors followed a comprehensive training course and assessments followed the meticulously developed and validated AQTT[35]. Furthermore, we attempted to mask the audio-recordings for information about organisation and

triage professional. Also, assessors were not aware of the design with both randomly selected and high-risk calls. Moreover, we dichotomised ratings of accuracy of triage, which was supported by the satisfactory inter-rater agreement of the AQTT [35]. Another limitation was our small sample size. We used data from a larger-scale study, and as a smaller proportion of calls was assessed as *clinically relevant* under- and overtriage, the study sample was not designed to find statistically significant differences. Therefore, we aimed to explore clinically relevant associations of patient- and call characteristics with inaccurate triage. Moreover, potential confounding could be an issue and the study was underpowered to conduct multivariate adjusted analyses to test for confounders. Finally, we needed to exclude calls for which the level of accuracy was assessed as not appicable, as it could both reflect a correct performance (i.e., "*this item correctly found not relevant")* or potentially cover a *poor* performance (i.e., "*available information is insufficient for assessment*").

*Interpretation and comparison of results*

Our study revealed that profession of triage staff and time of call had an effect on accuracy of triage. In a prior study, we found that nurse-led triage had less *clinically relevant* undertriage and more *clinically relevant* overtriage than GP-led triage in randomly selected calls[23]. In this study, we found the same tendencies in a selection of potential high-risk calls. Several factors could be related to the difference between nurse-led and GP-led triage, such as the role of CDSS, professional background, working- and organisational conditions. Nurses at the MH-1813 were obligated to use CDSS. CDSSs aim to ensure consistency[36,37] and to have a high degree of safety (i.e., low level of undertriage), which consequently leads to a higher level of overtriage. Telephone triage in OOH primary care serves as a sole entry point to health care in Denmark, being a gatekeeper of the healthcare system. A qualitative study described that nurses did not consider themselves as "gatekeepers", but as "service providers"[38]. Hence, perceptions of the task at hand may differ between nurses and GPs, thereby affecting overtriage, which could be seen as a service to the callers.

We found that high-risk calls during nighttime were significantly more likely to be undertriaged than high-risk calls during day or evening. This corresponds to a study by Hayward et al finding that patients calling during low call volume (e.g., nighttime) had a higher risk of requiring secondary care within three days after the OOH-PC contact[17]. Moreover, in paediatric patients calls conducted during nighttime were more likely to be defined as potential undertriaged[22]. Our study cannot elicit the mechanisms behind the increased risk of undertriage during nighttime. A stricter triage and gatekeeping function may be conducted during nighttime due to different organisational setup with less capacity of consultations and personal. Moreover, the threshold for patients to call OOH-PC may be different during nighttime and triage of the potentially more acutely ill patients may be more challenging.

We found no significant association between age, patient sex or weekday and the risk of undertriage or overtriage. However, our study may suggest that being elderly could influence the risk of undertriage, as we found a non-significantly trend of more undertriage in elderly compared to younger adults in high-risk calls. This corresponds to prior studies suggesting that increasing age was associated with unsafe triage[14,17,21].

*Implications for future research and clinical practice*

Although the risk of inaccurate triage differed between nurses and GPs, knowledge of mechanisms behind this difference are lacking and need further exploration. Our study found that calling during nighttime is associated with higher risk of being undertriaged, which may be related to a change in available resources and an urge to increase gate keeping. Thus, future qualitative studies should investigate the influence of availability of resources and organisational setup on the accuracy of telephone triage. Moreover, the influence of other patient characteristics (i.e. socioeconomically factors and age) and of health complaints presented in the call are relevant themes to study in relation to the accuracy of triage. From a clinical perspective this study suggests that triage professionals preferably should pay extra attention when

determining the appropriate triage decision in calls concerning abdominal pain conducted during nighttime and that extra attention may be focused when call is concerning elderly.

**CONCLUSION**

This study found that high-risk calls triaged by nurses were less likely to be undertriage and more likely to be overtriage compared to calls triaged by GPs in OOH-PC. In a Danish setting, the risk of under- and overtriage was similar in a randomly selected population and a high-risk patient population. High-risk calls conducted during nighttime were significantly more likely to be undertriaged than those during day and evening.

**Author statement**

All authors contributed to the development of the study protocol and design. DSG collected the calls, conducted the statistical analyses, and produced the first draft of the manuscript. DSG, AFP, LH, MBC, and FF contributed to the interpretation of data and critically revised the manuscript. All authors contributed with proofreading of the manuscript.

**Acknowledgements**

The authors would like to thank the 24 assessors who participated in the assessment process and the patients who gave valuable feedback during the development of the AQTT. The authors thank the MH-1813 and the GPC organisations. A special acknowledgment to statistician Claus H Vestergaard for providing valuable feedback in analyses.

**Competing interests**

None declared.

**Ethics approval**

The National Committee on Health Research Ethics in the Central Denmark Region was consulted and found that no approval was required for this study.

**Patient consent**

Consent was not obtained as approved by the Danish Patient Safety Authority (reference number: 3-3013-1274/1). The project (ID: 200) has been approved and is registered in the Record of Processing Activities at the Research Unit for General Practice in Aarhus in accordance with the provisions of the General Data Protection Regulation (GDPR).

**Transparency statement**

The lead author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as originally planned (and, if relevant, registered) have been explained.

**Funding**

This study was supported by the Danish foundation TrygFonden, Primary Health Care Research Foundation of the Central Denmark Region (*Praksisforskningsfonden*), the Committee for Quality Improvement and Continuing Medical Education in general practice in the Central Denmark Region (*Kvalitets- og Efteruddannelsesudvalget*) and the Committee of Multipractice Studies in General Practice (Multipraksisudvalget). Researchers were independent from funders.

**Data sharing statement**

Anonymised data are available upon reasonable request.

## References

1. Smits M, Rutten M, Keizer E, Wensing M, Westert G, Giesen P. The Development and Performance of After-Hours Primary Care in the Netherlands: A Narrative Review. Ann Intern Med. 2017;166:737–42.

2. Christensen MB, Olesen F. Out of hours service in Denmark: evaluation five years after reform. Bmj. 1998;316:1502–6.

3. Anderson A, Roland M. Potential for advice from doctors to reduce the number of patients referred to emergency departments by NHS 111 call handlers: Observational study. BMJ Open. 2015;5:1–4.

4. Bunn F, Byrne G, Kendall S. Telephone consultation and triage: effects on health care use and patient satisfaction. Cochrane Database of Systematic Reviews. 2004;

5. Salk ED, Schriger DI, Hubbell KA, Schwartz BL. Effect of visual cues, vital signs, and protocols on triage: A prospective randomized crossover trial. Annals of Emergency Medicine. 1998;32:655–64.

6. Morony S, Weir K, Duncan G, Biggs J, Nutbeam D, Mccaffery K. Enhancing communication skills for telehealth: development and implementation of a Teach-Back intervention for a national maternal and child health helpline in Australia. BMC Health Serv Res. 2018;18:162.

7. Agha Z, Roter D, Schapira RM. An evaluation of patient-physician communication style during telemedicine consultations. J Med Internet Res. 2009;11.

8. Ernesäter A, Winblad U, Engström M, Holmström IK. Malpractice claims regarding calls to Swedish telephone advice nursing: what went wrong and why? Journal of Telemedicine and Telecare. 2012;18:379–83.

9. Huibers L, Smits M, Renaud V, Giesen P, Wensing M. Safety of telephone triage in out-of-hours care: A systematic review. Scandinavian Journal of Primary Health Care. 2011;29:198–209.

10. Morreel S, Colliers A, Remmen R, Verhoeven V, Philips H. How accurate is telephone triage in out-of-hours care? An observational trial in real patients How accurate is telephone triage in out-of-hours care? An observational trial in real patients. Acta Clinica Belgica International Journal of Clinical and Laboratory Medicine. 2020;

11. Hansen EH, Hunskaar S. Telephone triage by nurses in primary care out-of-hours services in Norway: An evaluation study based on written case scenarios. BMJ Quality and Safety. 2011;20:390–6.

12. Smits M, Hanssen S, Huibers L, Giesen P. Telephone triage in general practices: A written case scenario study in the Netherlands. Scand J Prim Health Care. 2016;34:28–36.

13. Montalto M, Dunt DR, Day SE, Kelaher MA. Testing the safety of after-hours telephone triage: Patient simulations with validated scenarios. Australasian Emergency Nursing Journal. 2010;13:7–16.

14. Rørtveit S, Meland E, Hunskaar S. Changes of triage by GPs during the course of prehospital emergency situations in a Norwegian rural community. Scand J Trauma Resusc Emerg Med. 2014;21:89.

15. Kristoffersen JE. Out-of-hours primary care and the patients who die. A survey of deaths after contact with a suburban primary care out-of-hours service. Scandinavian Journal Of Primary Health Care. 2000;18:139–42.

16. Philips H, van Bergen J, Huibers L, Colliers A, Bartholomeeusen S, Coenen S, et al. Agreement on urgency assessment between secretaries and general practitioners: an observational study in out-of-hours general practice service in Belgium. Acta Clinica Belgica. 2015;3286:1–6.

17. Hayward GN, Vincent C, Lasserson DS. Predicting clinical deterioration after initial assessment in out-of-hours primary care: a retrospective service evaluation. British Journal of General Practice. 2016;1–8.

18. Gamst-Jensen H, Lippert FK, Egerod I. Under-triage in telephone consultation is related to non-normative symptom description and interpersonal communication: a mixed methods study. Scand J Trauma Resusc Emerg Med. 2017;25:1–8.

19. Elliott AM, McAteer A, Heaney D, Ritchie LD, Hannaford PC. Examining the role of Scotland's telephone advice service (NHS 24) for managing health in the community: analysis of routinely collected NHS 24 data. BMJ Open. 2015;5:e007293.

20. North F, Odunukan O, Varkey P. The value of telephone triage for patients with appendicitis. Journal of Telemedicine and Telecare. 2011;17:417–20.

21. Smits M, Huibers L, Kerssemeijer B, de Feijter E, Wensing M, Giesen P. Patient safety in out-of-hours primary care: a review of patient records. BMC Health Serv Res. 2010;10:1–9.

22. Kempe A, Bunik M, Ellis J, Magid D, Hegarty T, Dickinson LM, et al. How Safe Is Triage by an After-Hours Telephone Call Center? Pediatrics. 2006;118:457–63.

23. Graversen DS, Christensen MB, Pedersen AF, Carlsen AH, Bro F, Christensen HC, et al. Safety, efficiency and health-related quality of telephone triage conducted by general practitioners, nurses, or physicians in out-of-hours primary care: A quasi-experimental study using the Assessment of Quality in Telephone Triage (AQTT) to assess audio. BMC Family Practice. 2020;21:1–12.

24. Ernesäter A, Holmström IK, Engström M. Telenurses' experiences of working with computerized decision support: Supporting, inhibiting and quality improving. Journal of Advanced Nursing. 2009;65:1074–83.

25. Holmström IK. Decision aid software programs in telenursing: Not used as intended? Experiences of Swedish telenurses. Nursing and Health Sciences. 2007;9:23–8.

26. Holmström IK, Kaminsky E, Lindberg Y, Douglas Spangler |, Winblad U. Registered Nurses' experiences of using a clinical decision support system for triage of emergency calls: A qualitative interview study. 2020;

27. Kempe A, Dempsey C, Whitefield J, Bothner J, MacKenzie T, Poole S. Appropriateness of urgent referrals by nurses at a hospital-based pediatric call center. Archives of Pediatrics and Adolescent Medicine. 2000;154:355–60.

28. Belman S, Murphy J, Steiner JF, Kempe A. Consistency of triage decisions by call center nurses. Ambul Pediatr. 2002;2:396–400.

29. Pasini A, Rigon G, Vaona A. A cross-sectional study of the quality of telephone triage in a primary care out-of-hours service. Journal of Telemedicine and Telecare. 2015;21:68–72.

30. Woodring BC. Pediatric telephone triage protocols: standardized decision-making or a false sense of security? J Child Fam Nurs. 2000;3:60–2.

31. Staub GM, von Overbeck J, Blozik E. Teleconsultation in children with abdominal pain: a comparison of physician triage recommendations and an established paediatric telephone triage protocol / La teleconsulta en niños con dolor abdominal: una comparación de recomendaciones de triage médico . BMC Medical Informatics and Decision Making. 2013;13:110.

32. Graversen DS, Huibers L, Christensen MB, Bro F, Collatz Christensen H, Vestergaard CH, et al. Communication quality in telephone triage conducted by general practitioners, nurses or physicians: A quasi-experimental study using the AQTT to assess audio-recorded telephone calls to out-of-hours primary care in Denmark. BMJ Open. 2020;10:1–12.

33. Statistics Denmark. StatBank Denmark [Internet].

34. Ebert JF, Huibers L, Lippert FK, Christensen B, Christensen MB. Development and evaluation of an "emergency access button" in Danish out-of-hours primary care: a study protocol of a randomized controlled trial. BMC Health Services Research. 2017;17:1–8.

35. Graversen DS, Pedersen AF, Carlsen AH, Bro F, Huibers L, Christensen MB. Quality of out-of-hours telephone triage by general practitioners and nurses: development and testing of the AQTT - an assessment tool measuring communication, patient safety and efficiency. Scandinavian Journal Of Primary Health Care. 2019;37:18–29.

36. Murdoch J, Barnes R, Pooler J, Lattimer V, Fletcher E, Campbell JL. The impact of using computer decision-support software in primary care nurse-led telephone triage: interactional dilemmas and conversational consequences. Social Science & Medicine. 2015;126:36–47.

37. LT W, Zwart D, Erkelens DC, Huijsmans M, Hoes AW, Damoiseaux R, et al. Tinkering and overruling the computer decision support system: Working strategies of telephone triage nurses who assess the urgency of callers suspected of having an acute cardiac event. Journal of Clinical Nursing. 2019;1175–86.

38. Lindberg BH, Rebnord IK, Høye S. Phone triage nurses' assessment of respiratory tract infections–the tightrope walk between gatekeeping and service providing. A qualitative study. Scandinavian Journal of Primary Health Care. 2021;39:139–47.

**Figure and Table Legends**

**Box 1.** Description of the organisation and telephone triage of the two included models for out-of-hours primary care

**Figure 1.** Flowchart of inclusion and exclusion of randomly selected and high-risk calls from the GPC and MH-1813

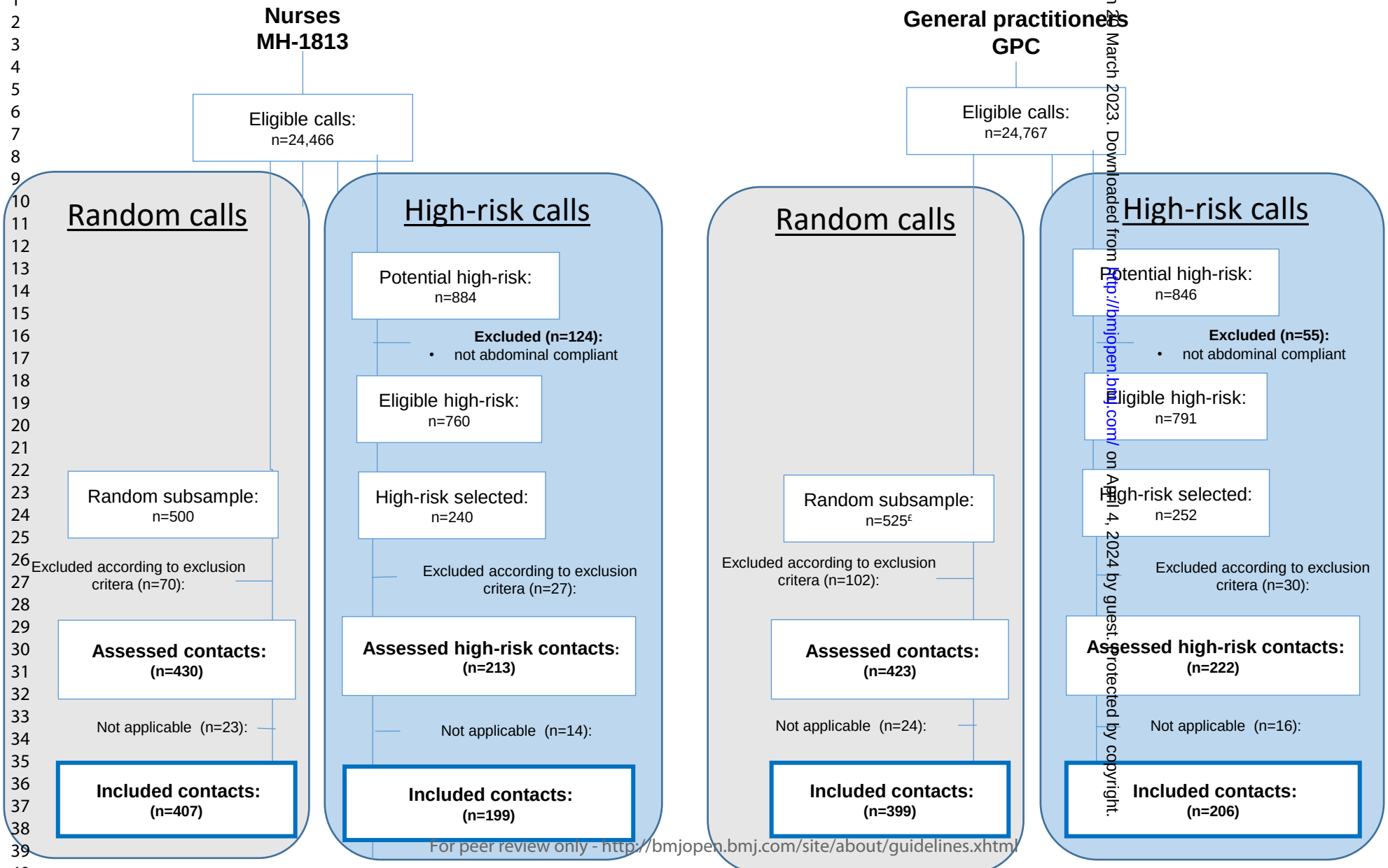
*Legends: Not applicable:* accuracy of triage could not be assessed due to insufficient information in the call or assessment of accuracy of triage was deemed not relevant.

**Table 1.** Baseline description of patient and call characteristics in randomly selected calls and high-risk calls

**Table 2.** Patient- and call characteristics associated with clinically relevant under- and overtriage in randomly selected calls[1]

**Table 3.** Patient- and call characteristics associated with clinically relevant under- and overtriage in high-risk calls

**Appendix 1.** Exclusion criteria



**Appendix 1.** Exclusion criteria

| Type of call | Definition/clarification |
|---|---|
| Frequent callers | Patients with ≥7 calls during the two-week inclusion period (the patient's medical record from the OOH service that could include important information on these patients was only available to the triage professional and not to the assessor) |
| Calls by mistake | Calls with no caller answering the triage professional |
| Daytime calls | Calls performed during daytime (the telephone triage service at MH-1813 was available during daytime) |
| Calls by other health professionals | The caller was another healthcare professional, e.g. from a nursing home |
| Administrative calls | The reason for calling was administrative, e.g. calling to get the number for the acute dentist |
| Calls regarding simple drug prescriptions | The patient called for renewal of a prescription that required little information sharing |
| Preterm termination | Calls that were ended too early, e.g. calls made by error, no sound on call, or sound interrupted in the middle of call |
| Calls from other localisation | Calls from a caller who was not in the same location as the patient, e.g. parent on the way to pick up a sick child from day care |
| Calls with poor sound quality | Calls with poor sound quality making assessment difficult |
| Language issues | Calls in which language issues challenged the triage, i.e. caller did not speak Danish or English |

| | |
|---|---|
| Not able to identify call | Calls where an exact linkage to the corresponding audio recording could not be established |

# Reporting checklist for cross sectional study.

Based on the STROBE cross sectional guidelines.

## Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the STROBE cross sectionalreporting guidelines, and cite them as:

von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies.

| | | Reporting Item | Page Number |
|---|---|---|---|
| **Title and abstract** | | | |
| Title | #1a | Indicate the study's design with a commonly used term in the title or the abstract | 1 |

| | | | |
|---|---|---|---|
| Abstract | #1b | Provide in the abstract an informative and balanced summary of what was done and what was found | 3 |
| **Introduction** | | | |
| Background / rationale | #2 | Explain the scientific background and rationale for the investigation being reported | 5 |
| Objectives | #3 | State specific objectives, including any prespecified hypotheses | 5 |
| **Methods** | | | |
| Study design | #4 | Present key elements of study design early in the paper | 5-6 |
| Setting | #5 | Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection | 5-8 |
| Eligibility criteria | #6a | Give the eligibility criteria, and the sources and methods of selection of participants. | 6-7 |
| | #7 | Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable | 7-8 |
| Data sources / measurement | #8 | For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Give information separately for for exposed and unexposed groups if applicable. | 7-8 |

| | | | |
|---|---|---|---|
| Bias | #9 | Describe any efforts to address potential sources of bias | 7-8 |
| Study size | #10 | Explain how the study size was arrived at | 6-7 |
| Quantitative variables | #11 | Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why | 8-9 |
| Statistical methods | #12a | Describe all statistical methods, including those used to control for confounding | 8-9 |
| Statistical methods | #12b | Describe any methods used to examine subgroups and interactions | 8-9 |
| Statistical methods | #12c | Explain how missing data were addressed | 9 |
| Statistical methods | #12d | If applicable, describe analytical methods taking account of sampling strategy | 9 |
| Statistical methods | #12e | Describe any sensitivity analyses | 9 |
| **Results** | | | |
| Participants | #13a | Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed. Give information separately for for exposed and unexposed groups if applicable. | 6-7 |

| | | | |
|---|---|---|---|
| Participants | #13b | Give reasons for non-participation at each stage | 6-8, appendix, flow diagram |
| Participants | #13c | Consider use of a flow diagram | 6-8, flow diagram |
| Descriptive data | #14a | Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Give information separately for exposed and unexposed groups if applicable. | 6-7,9, table 1 |
| Descriptive data | #14b | Indicate number of participants with missing data for each variable of interest | 9 |
| Outcome data | #15 | Report numbers of outcome events or summary measures. Give information separately for exposed and unexposed groups if applicable. | 9-10, Table 1,2 and 3 |
| Main results | #16a | Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included | 9-10, Table 1,2,3 |
| Main results | #16b | Report category boundaries when continuous variables were categorized | 8 |

| | | | |
|---|---|---|---|
| Main results | #16c | If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period | 9-10 |
| Other analyses | #17 | Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses | 9-10 |
| **Discussion** | | | |
| Key results | #18 | Summarise key results with reference to study objectives | 11 |
| Limitations | #19 | Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias. | 11-12 |
| Interpretation | #20 | Give a cautious overall interpretation considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence. | 12-13 |
| Generalisability | #21 | Discuss the generalisability (external validity) of the study results | 13 |
| **Other Information** | | | |
| Funding | #22 | Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based | 15 |

Notes:

- 13b: 6-8, appendix, flow diagram
- 13c: 6-8, flow diagram

- 14a: 6-7,9, table 1
- 15: 9-10, Table 1,2 and 3
- 16a: 9-10, Table 1,2,3  This checklist was completed on 19. May 2022 using https://www.goodreports.org/, a tool made by the EQUATOR Network in collaboration with Penelope.ai

BMJ Open

# Factors associated with under- and overtriage in telephone triage in Danish out-of-hours primary care – a natural quasi-experimental cross-sectional study of randomly selected and high-risk calls

| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2022-064999.R1 |
| Article Type: | Original research |
| Date Submitted by the Author: | 22-Dec-2022 |
| Complete List of Authors: | Graversen, Dennis; Aarhus University Research Unit General Practice; Aarhus University, Department of Public Health<br>Pedersen, Anette; Aarhus University Research Unit General Practice; Aarhus University, Department of clinical medicine<br>Christensen, Morten; Aarhus University Research Unit General Practice; Aarhus University, Department of Public Health<br>Folke, Fredrik; Copenhagen University Hospital, Emergency Medical Services Copenhagen; Copenhagen University Hospital, Department of Cardiology - Herlev and Gentofte<br>Huibers, L; Aarhus University Research Unit General Practice, Research Unit for General Practice |
| Primary Subject Heading: | General practice / Family practice |
| Secondary Subject Heading: | Nursing |
| Keywords: | Clinical audit < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health & safety < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, PRIMARY CARE |
| | |

*I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.*

*The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.*

*Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.*

# Factors associated with under- and overtriage in telephone triage in Danish out-of-hours primary care – a natural quasi-experimental cross-sectional study of randomly selected and high-risk calls

**Graversen DS[1,2], Pedersen AF[1,3], Christensen MB[1,2], Folke F[4,5,6], Huibers L[1]**

[1]Research Unit for General Practice, Aarhus, Denmark, [2]Department of Public Health, Aarhus University, Aarhus, Denmark, [3]Department of Clinical Medicine, Aarhus University, Aarhus, Denmark, [4]Copenhagen University Hospital - Emergency Medical Services Copenhagen, Ballerup, Denmark, [5]Department of Cardiology, Copenhagen University Hospital - Herlev and Gentofte, Denmark, [6]Department of Clinical Medicine, University of Copenhagen, Denmark

*Correspondence to*

DS Graversen (ORCID ID: 0000-0002-7784-9403), Research Unit for General Practice, Aarhus, Bartholins Allé 2, 8000 Aarhus C, Denmark & Department of Public Health, Aarhus University, Bartholins Allé 2, 8000 Aarhus C, Denmark

E-mail: d.graversen@ph.au.dk, telephone: +45 28 26 05 52

Keywords: Triage, Telephone, After-Hours Care, Primary Health Care, General Practitioners, Nurses, Safety, Efficiency, Quality of Health Care

Abstract: words 284

Word count: 2953

**Abbreviations:**

AQTT: Assessment of Quality in Telephone Triage (a validated tool for assessment of triage performance)

CDSS: Computerised decision support system

GP: General practitioners

GPC: General practitioners cooperative

MH-1813: Medical helpline 1813

OOH: Out-of-hours

OOH-PC: Out-of-hours primary care

**ABSTRACT:**

**Objectives**: We aim to explore under- and overtriage in a high-risk patient population and explore patient- and call characteristics associated with under- and overtriage in both randomly selected and in high-risk telephone calls to out-of-hours primary care (OOH-PC).

**Design:** Natural quasi-experimental cross-sectional study.

**Setting:** Two Danish OOH-PC services using different telephone triage models: a general practitioner cooperative (GPC) with GP-led triage and the medical helpline 1813 with CDSS guided nurse-led triage.

**Participants:** We included audio-recorded telephone triage calls from 2016: 806 random calls and 405 high-risk calls (defined as patients ≥ 30 years calling with abdominal pain).

**Main outcome measures:** Twenty-four experienced physicians used a validated assessment tool to assess the accuracy of triage (AQTT). We calculated the relative risk (RR) for *clinically relevant* under- and overtriage for a range of patient- and call characteristics.

**Results:** We included 806 randomly selected calls (44 *clinically relevant* undertriaged and 54 *clinically relevant* overtriaged) and 405 high-risk calls (32 undertriaged and 24 overtriaged). In high-risk calls, nurse-led triage was associated with significantly less undertriage (RR: 0.47, 95%CI: 0.23-0.97) and more overtriage (RR: 3.93, 95%CI: 1.50-10.33) compared to GP-led triage. In high-risk calls, the risk of undertriage was significantly higher for calls during nighttime (RR: 2.1, 95%CI: 1.05-4.07). Undertriage tended to be more likely for calls concerning patients ≥ 60 years compared to 30 to 59 years (11.3% vs. 6.3%) in high-risk calls. However, this result was not significant.

**Conclusion:** Nurse-led triage was associated with less undertriage and more overtriage compared to GP-led triage in high-risk calls. This study may suggest that to minimize undertriage, the triage professionals should pay extra attention when a call occurs during nighttime or concerns elderly. However, this needs confirmation in future studies.

**Strengths and limitations of this study**

- A strength of our study is the natural quasi-experimental design using real-life calls as opposed to a constructed setup.
- To our knowledge, this is the first study to explore factors associated with undertriage and overtriage in real audio-recorded calls.
- A limitation is the use of only one assessor per call, but acceptable interrater agreement was previously found.

## INTRODUCTION

In out-of-hours primary care (OOH-PC), telephone triage plays a pivotal role in managing patient flows and workload[1–3]. Telephone triage aims to ensure a safe and efficient delivery of healthcare, avoiding undertriage and minimizing overtriage[4]. However, accurate telephone triage is difficult due to the lack of visual cues of the patient, challenges in telecommunication, and time pressure[5–8].

Safety and efficiency of telephone triage in OOH-PC have been explored in a range of studies that used varying outcome measures[4,9–13]. Studies have identified a range of risk factors for potentially unsafe telephone triage in out-of-hours (OOH) care: calls concerning abdominal pain[14–20], chest pain[16,17] and shortness of breath[16,18], calls for patients with increasing age[14,17,21], and calls during nighttime[17,22]. Thus, these calls can be seen as potential high-risk calls. A few studies have explored factors associated with overtriage as a measure of inefficient telephone triage. Nurse-led triage [1,23] and triage aided by computerized decision support system (CDSS)[23–26] have been associated with overtriage. However, studies that explored factors associated with under- and overtriage used a range of study designs. None of these studies used real audio-recorded telephone calls to assess the risk of under- and overtriage and related risk factors.

Previously, we found that telephone triage was associated with significantly less undertriage and significantly more overtriage for nurse-led triage compared to GP-led triage in a sample of random calls[23] . However, it remains unclear whether this association also exists for high-risk calls. Therefore, we aim to investigate the risk of under- and overtriage in high-risk telephone calls to OOH-PC in Denmark. In addition, we aim to explore patient- and call characteristics associated with under- and overtriage in random calls and in high-risk calls to OOH-PC.

## METHODS

*Design and setting*

This paper presents secondary analyses of a natural quasi-experimental study in two regional OOH-PC services in Denmark[23,27]. In one prior paper [23], we explored safety and efficiency in randomly selected calls, using under- and overtriage. In the present paper, we included potentially high-risk calls and studied under- and overtriage of these calls. Subsequently, we explored factors associated with under- and overtriage in randomly selected calls and in high-risk patient calls.

OOH-PC is provided by the general practitioner cooperative (GPC) in the Central Denmark Region since 1992 and by the medical helpline 1813 (MH-1813) in the Capital Region of Denmark since 2014. Both services are open outside office hours (i.e., on weekdays from 4 pm to 8 am, weekends, and national holidays), offering telephone consultations, clinic consultations, and home visits. The GPC and MH-1813 use different triage models. The GPC uses GP-led telephone triage without CDSS, whereas the MH-1813 uses nurse-led telephone triage (see Box 1). At the MH-1813, the majority of incoming unselected calls (i.e., 74%) are answered by registered nurses[28]. Nurses are obliged to use a locally developed CDSS and they can redirect calls to a physician on call (i.e., 11% of all incoming unselected calls answered by a nurse)[28].

**Box 1.** Description of the organisation and telephone triage of the two included models for out-of-hours primary care

| | **GP cooperative (GPC)** | **Medical helpline 1813 (MH-1813)** |
|---|---|---|
| **Region** | **Central Denmark Region** | **Capital Region of Denmark** |
| **Population**[29] | 1.3m citizens | 1.8m citizens |
| **Telephone calls in 2016**[28] | 701,000 | 806,000 |

| | | |
|---|---|---|
| **Organiser** | GPs in the region | Regional administration |
| **Organisation and services** | ▪ Telephone consultations, home visits, and clinic consultations at the GPC<br>▪ GPs are obliged to take part in the service | ▪ Telephone consultations and home visits run by MH-1813<br>▪ Clinic consultations are located in hospital facilities and managed by EDs |
| **Remuneration of professionals** | Fee for service | Payment by the hour |
| **Triage model** | GPs or GP trainees in their final year of speciality; no CDSS available<br><br>GPs typically work 8-hour shifts and have approx. 1-4 shifts per month besides their daytime work | Nurses who are obliged to use a CDSS and have option to redirect calls to a physician<br>Physicians with different medical specialities (a minority being GPs)<br>Nurses work 8-hour shifts and are mostly fully employed at MH-1813.<br>Physicians work 8-hour shifts and are employed besides their daytime job |

*CDSS: Computerized decision support system; The triage models and organisations remain largely unchanged since 2014.*

*Definition of groups of calls*

We defined two groups of calls: 1) a group of randomly selected calls that was representative for all calls to OOH-PC and 2) a group of potentially high-risk patient calls (referred to as "high-risk"). To define the high-risk patient calls, we conducted a systematic literature search in 2016 to identify factors associated with unsafe or inaccurate telephone triage. Seven identified studies described factors associated with a risk of

unsafe telephone triage: infants[22], increasing patient age[14,17,21], calls during nighttime[17,22], abdominal pain[14–18], chest pain[16,17], and breathing difficulties[16,18]. In a consensus meeting, the authors (DSG – medical doctor, AFP - psychologist, MBC - general practitioner, LH - medical doctor) defined criteria for high-risk calls as calls concerning patients above 30 years who suffered from abdominal pain. We added the age criterion, as we aimed to include potentially dangerous conditions that present with vague, indistinctly, or greatly differing symptomatology (such as dissecting aorta aneurism, myocardial infarction, cholecystitis, pyelonephritis, acute pancreatitis, gastrointestinal ulcer, and gynaecological causes). Due to pragmatic considerations, we focused on one reason for encounter; selecting high-risk calls had to be done by manual development of an algorithm (see further).

*Selection of calls*

All calls to the GPC and MH-1813 outside office hours during the inclusion period (GPC: 23 November - 7 December 2016; MH-1813: 23 November - 8 December 2016) were potentially eligible. For calls redirected by a nurse to a physician at MH-1813, only the part conducted by the nurse was eligible and later assessed as described below. Randomly selected calls were selected from all eligible calls; these calls were the same as we studied in a prior paper [23]. The aim was to include 435 calls by GPs and 435 by nurses, based on a power calculation to detect a 5% difference in undertriage between nurse-led and GP-led undertriage in line with the aim of the main study[23]. Based on expected exclusion, we randomly selected 525 random calls from GPC and 500 calls from MH-1813.

To identify potential eligible high-risk calls, we received electronic patient journal records from the GPC and MH-1813. We defined a list of wordings and abbreviations associated with abdominal pain, which was used to search the patient records and identify all eligible calls concerning abdominal pain. DSG (medical doctor) assessed these marked patient records to check whether inclusion for abdominal pain was correct. Calls were excluded when the triage professional in the patient record dismissed the presence of abdominal pain

(i.e., triagist noted *no abdominal pain*) or when the complaint was clearly outside the thoracic, abdominal, or pelvic region. Thus, we identified 846 (GPC) and 884 (MH-1813) eligible high-risk calls for potential inclusion. A power calculation was designed to detect a significant difference between nurse-led and GP-led undertriage for high-risk calls, without detecting associations between risk factors and over-and undertriage. As this calculation revealed the need of 206 calls per triage model, we randomly selected 252 calls from the GPC and 240 calls from MH-1813. Selection of both randomly selected and high-risk calls was done, matching the overall distribution on day of week (i.e., weekend, not weekend) and time of day (i.e., day, evening, night).

Each selected call had a unique identification number that was used to identify the corresponding audio-recording. Three master students of medicine masked the audio-recordings using beep tones to cover information revealing triage profession, OOH organisation, and patient identification information. In addition, the students screened all calls for fulfilling exclusion criteria as shown in Appendix 1. DSG reviewed all calls that fulfilled exclusion criteria and those that were unclear. In case of doubt, DSG and AFP reached consensus regarding exclusion.

*Assessment of accuracy of triage*

All included calls have been assessed as described in prior papers,[23,30] using the tool "Assessment of Quality in Telephone Triage" (AQTT). The AQTT comprises 24 items assessing the health-related quality and the quality of communication[30]. AQTT has a rating manual, describing when to assign the different ratings for each item[30]. In the present study, the outcome accuracy of triage was measured by one item that used a 7-point scale to differentiate between levels of undertriage (ratings 1 to 3), optimal triage (rating 4), and overtriage (ratings 5 to 7). The AQTT showed inter-rater disagreement when using the entire 7-point scale for assessment of the accuracy of triage, but revealed satisfactory inter-rater agreement when

using dichotomous scales for *clinically relevant* undertriage (ratings "1" and "2" versus ratings "3" to "7") and *clinically relevant* overtriage (ratings "6" and "7" versus ratings "1" to "5")[30].

For calls that triage nurses redirected to a physician, only the part conducted by the nurse was available for assessment. These calls could be assessed as optimal, if the decision to redirect the call was what would be expected by a nurse.

For the assessment panel, we recruited 24 physicians among triage professionals from the GPC and MH-1813 using two inclusion criteria: 1) >1 year experience and 2) active in telephone triage in OOH-PC at time of study. We randomly selected 16 GPs from the 56 interested GPs from the GPC. At MH-1813, we included all eight physicians fulfilling our inclusion criteria from the ten interested physicians. All assessors followed a two-day training course providing knowledge on telephone triage and communication, introducing the AQTT and rating manual, and assessing triage calls individually and in plenary, focusing on achieving consistency. We randomly distributed calls to all assessors, so each member of the assessment panel assessed random and high-risk calls triaged with both the GP-led model and the nurse-led model. Assessors were blinded for the type of call and triage model. Information on age and sex of the patient, day of week, and the time of each call was available for the assessor, extracted from the registration systems from the GPC and MH-1813.

*Statistical analyses*

Accuracy of triage decision was categorised into *clinically relevant* undertriage (rated "1" or "2") and *clinically relevant* overtriage (rated "6" or "7"). We conducted separate analyses for randomly selected calls and for high-risk calls. We used descriptive analyses to describe patient- and call characteristics, as well as the risk of *clinically relevant* undertriage and *clinically relevant* overtriage for both type of calls. We explored the association of individual patient characteristics (i.e., age, patient sex) and call characteristics

(i.e., weekend, time of day, triage model) with inaccurate telephone triage, by calculating the risk ratio (RR) of having *clinically relevant* undertriage (vs. no *clinically relevant* undertriage) and *clinically relevant* overtriage (vs. no *clinically relevant* overtriage), using binomial regression. 95% confidence intervals (95%CI) were calculated. All analyses were performed in Stata 14.2 (StataCorp. 2015. *Stata Statistical Software: Release 14.2*. College Station, TX: StataCorp LP).

**PUBLIC AND PATIENT INVOLVEMENT STATEMENT**

We explored patients' perspective in a focus group interview concerning the development of the AQTT and incorporated received input in the rating manual of the AQTT [30].

**RESULTS**

*Description of calls*

Figure 1 shows the flowchart of selection and exclusion of calls. We excluded 47 randomly selected calls and 30 high-risk calls assigned 'not applicable', as assessing accuracy of triage was not possible (e.g., insufficient information was available) or not relevant. Thus, our final study population included 806 randomly selected calls and 405 high-risk calls (Table 1). In high-risk calls, the risk of *clinically relevant* undertriage was 7.9%, whereas the risk of *clinically relevant* overtriage was 5.9%.

*(Figure 1 – flowchart- attached separate)*

**Figure 1.** Flowchart of inclusion and exclusion of randomly selected and high-risk calls from the GPC and MH-1813

*Not applicable:* accuracy of triage could not be assessed due to insufficient information in the call or assessment of accuracy of triage was deemed not relevant.

**Table 1.** Baseline description of patient and call characteristics in randomly selected calls and high-risk calls

| | **Randomly selected calls** | **High-risk calls[1]** |
|---|---|---|
| | **N=806[2]** | **N=405[2]** |
| ***Patient characteristics*** | % (N) | % (N) |
| **Age groups (in years)** | | |
| - <18 | 36.5 (294) | N.a.[1] |
| - 18-29 | 21.0 (169) | N.a.[1] |
| - 30-59 | 27.5 (222) | 67.2 (272) |
| - ≥60 | 15.0 (121) | 32.8 (133) |
| **Sex** | | |
| - Male | 40.1 (323) | 40.3 (163) |
| - Female | 59.9 (483) | 59.8 (242) |
| ***Call characteristics*** | | |
| **Weekend** | | |
| - Not weekend | 46.4 (374) | 51.4 (208) |
| - Weekend | 53.6 (432) | 48.6 (197) |
| **Time of day** | | |
| - Day or evening | 83.6 (674) | 77.5 (314) |
| - Nighttime | 16.4 (132) | 22.5 (91) |
| **Assessed accuracy of triage[3]** | | |
| *Clinically relevant* undertriage | | |
| 1: Severe undertriage | 1.5 (12) | 3.0 (12) |
| 2: Moderate undertriage | 4.0 (32) | 4.9 (20) |

| | | |
|---|---|---|
| *Satisfactory* triage | | |
| 3: Mild undertriage | 8.6 (69) | 11.6 (47) |
| 4: Optimal triage | 69.0 (556) | 66.2 (268) |
| 5: Mild overtriage | 10.3 (83) | 8.4 (34) |
| *Clinically relevant* overtriage | | |
| 6: Moderate overtriage | 4.6 (37) | 4.4 (18) |
| 7: Severe overtriage | 2.1 (17) | 1.5 (6) |
| **Triage model** | | |
| - GP-led triage | 49.5 (399) | 50.9 (206) |
| - Nurse-led triage with CDSS | 50.5 (407) | 49.1 (199) |

[1]*High-risk calls only included patients aged ≥30 years calling OOH-PC with abdominal pain.* [2]*Not applicable: Calls where the assessment of accuracy of triage was not possible or not relevant were excluded (randomly selected: n=47, high-risk; n=30).* [3]*Accuracy of triage was categorized into three categories: 1) satisfactory triage (including optimal triage ("4"), mild undertriage ("3"), and mild overtriage ('5'), 2) clinically relevant undertriage ("1" or "2"), and 3) clinically relevant overtriage ("6" or "7").*
*Definitions: Weekend = Friday 4 PM to Monday 8 AM and holidays. Day = 8 AM to 4 PM, Evening = 4 PM to midnight, Night = midnight to 8 AM.*

*Risk factors for inaccurate triage in randomly selected calls*

For randomly selected calls, we found no significant association between patient characteristics (i.e., age and sex), call characteristics (i.e., weekend and time of call) and the risk of clinically relevant under- or overtriage (Table 2).

**Table 2.** Patient- and call characteristics associated with clinically relevant under- and overtriage in randomly selected calls[1]

| | **Undertriage[2]** | **Overtriage[3]** |
|---|---|---|
| | **N=44** | **N=54** |

| | % (N) | RR² (95%CI) | % (N) | RR² (95%CI) |
|---|---|---|---|---|
| ***Patient characteristics*** | | | | |
| **Age groups (in years)** | | | | |
| - <18 (n=294) | 3.7 (11) | 0.57 (0.23-1.37) | 7.5 (22) | 1.51 (0.63-3.63) |
| - 18-29 (n=169) | 4.1 (7) | 0.62 (0.23-1.68) | 7.7 (13) | 1.55 (0.61-3.97) |
| - 30-59 (n=222) | 8.1 (18) | 1.23 (0.55-2.74) | 5.9 (13) | 1.18 (0.46-3.03) |
| - ≥ 60 (n=121) | 6.6 (8) | 1 | 5.0 (6) | 1 |
| **Sex** | | | | |
| - Male (n=323) | 5.3 (17) | 1 | 4.6 (15) | 1 |
| - Female (n=483) | 5.6 (27) | 1.06 (0.59-1.92) | 8.1 (39) | 1.74 (0.97-3.10) |
| ***Call characteristics*** | | | | |
| **Weekend** | | | | |
| - Not weekend (n=374) | 5.4 (20) | 1 | 6.7 (25) | 1 |
| - Weekend (n=432) | 5.6 (24) | 1.04 (0.58-1.85) | 6.7 (29) | 1.00 (0.60-1.68) |
| **Time of day** | | | | |

| | | | | |
|---|---|---|---|---|
| - Day or evening (n=674) | 5.3 (36) | 1 | 7.0 (47) | 1 |
| - Nighttime (n=132) | 6.1 (8) | 1.13 (0.54-2.39) | 5.3 (7) | 0.76 (0.35-1.65) |

*[1]We excluded 47 calls, as assessing accuracy of triage was not possible or not relevant; [2]Undertriage: triage decision assessed as clinically relevant undertriage (rated "1" or "2"); [3]Overtriage: clinically relevant overtriage (rated "6" or "7") ; *Significant difference in RR, using binomial regression analyses; p<0.05: Definitions: Weekend = Friday 4 PM to Monday 8 AM and holidays. Day = 8 AM to 4 PM, Evening = 4 PM to midnight, Night = midnight to 8 AM.*



*Risk factors for inaccurate triage in high-risk calls*

Nurse-led triage was associated with significantly less *clinically relevant* undertriage than GP-led triage (RR=0.47, 95%CI: 0.23-0.97) in high-risk calls (Table 3). Nurse-led triage had significantly more *clinically relevant* overtriage (RR=3.93, 95%CI: 1.50-10.53) compared to GP-led triage. High-risk calls conducted during nighttime had a significantly higher risk of being *clinically relevant* undertriaged (13.2%) compared to calls conducted during day or evening (6.4%) (RR=2.1, 95%CI: 1.05-4.07). We found no significant association for the other patient- and call characteristics. However, a trend was seen for patient age, as the

risk of *clinically relevant* undertriage was lower in patients aged 30 to 59 years (6.3%) in comparison with elderly patients ≥60 years (11.3%) (RR=0.55, 95%CI: 0.29-1.07).

**Table 3.** Patient- and call characteristics associated with clinically relevant under- and overtriage in high-risk calls

| | **Undertriage[1]** | | **Overtriage[2]** | |
|---|---|---|---|---|
| | **N=32** | | **N=24** | |
| | % (N) | RR (95%CI) | % (N) | RR (95%CI) |
| ***Patient characteristics*** | | | | |
| **Age groups (in years)** | | | | |
| - 30-59 (n=272) | 6.3 (17) | 0.55 (0.29-1.07) | 7.4 (20) | 2.44 (0.85-7.01) |
| - >60 (n=133) | 11.3 (15) | 1 | 3.0 (4) | 1 |
| **Sex** | | | | |
| - Male (n=163) | 8.6 (14) | 1 | 6.8 (11) | 1 |
| - Female (n=242) | 7.4 (18) | 0.87 (0.44-1.69) | 5.4 (13) | 0.80 (0.40-1.73) |
| ***Call characteristics*:** | | | | |
| **Weekend** | | | | |
| - Not weekend (n=208) | 6.7 (14) | 1 | 6.3 (13) | 1 |
| - Weekend (n=197) | 9.1 (18) | 1.36 (0.69-2.65) | 5.6 (11) | 0.89 (0.41-1.95) |
| **Time of day** | | | | |

| | | | | |
|---|---|---|---|---|
| - Day or evening (n=314) | 6.4 (20) | 1 | 6.1 (19) | 1 |
| - Nighttime (n=91) | 13.2 (12) | 2.1 (1.05-4.07)* | 5.5 (5) | 0.91 (0.35-2.36) |
| **Triage model** | | | | |
| - GP-led triage (n=206) | 10.7 (22) | 1 | 2.4 (5) | 1 |
| - Nurse-led triage with CDSS (n=199) | 5.0 (10) | 0.47 (0.23-0.97)* | 10.0 (19) | 3.93 (1.50-10.33)* |

*We excluded 30 calls as assessing accuracy of triage was not possible or not relevant; [1]Undertriage: triage decision assessed as clinically relevant undertriage (rated "1" or "2"): [2]Overtriage: clinically relevant overtriage (rated "6" or "7"); *Significant difference in RR, using binomial regression analyses; p<0.05: Definitions: Weekend = Friday 4 PM to Monday 8 AM and holidays. Day = 8 AM to 4 PM, Evening = 4 PM to midnight, Night = midnight to 8 AM.*

## DISCUSSION

### *Principal findings*

In high-risk calls, nurse-led triage was associated with significantly lower risk of *clinically relevant* undertriage and significant higher risk of *clinically relevant* overtriage compared to GP-led triage. For high-risk calls, the risk of *clinically relevant* undertriage was significantly higher if the call was conducted during nighttime compared to day and evening. For randomly selected calls, we found no significant association between defined risk factors and the risk of *clinically relevant* under- or overtriage

### *Strengths and limitations of the study*

A major strength was the quasi-experimental design using real-life calls as opposed to the constructed setup used in previous studies [18–20, 31, 32]. An additional strength was the meticulous assessment process using the validated AQTT tool combined with a comprehensive rating manual [30].

Our study had the following limitations. Due to the thorough assessment process, each call was assessed by one assessor. Consequently, bias due to misclassification cannot be rejected. However, we took several precautions to ensure consistency of assessments. The assessors followed a comprehensive training course and assessments followed the meticulously developed and validated AQTT[30]. Furthermore, we attempted to mask the audio-recordings for information about organisation and triage model. Also, assessors were not aware of the design with both randomly selected and high-risk calls. Moreover, we dichotomised accuracy of triage into *clinically relevant* undertriage and overtriage, which had a satisfactory inter-rater agreement of the AQTT [30]. We carefully decided upon our definition of high-risk calls (i.e., calls for adults >30 years with abdominal pain). The cut-off point for age was reached through meticulous discussions among the authors, but could be too low, thereby including a larger group of low-risk calls. Another limitation was our small sample size. In line with the main study, our power calculation was made to identify a significant difference between nurse-led and GP-led under- and overtriage. As only a selection of these calls was assessed as *clinically relevant* under- or overtriage, the present study lacked power to identify risk factors for under- and overtriage. Therefore, we explored clinically relevant associations of patient- and call characteristics with inaccurate triage. Furthermore, we chose physicians as assessors of the accuracy of triage, as they were most frequently used in other studies. The decision to include only physicians in the assessment panel may have induced similar-to-me cognitive bias when assessing nurse-led triage, leading to underassessment of the accuracy of nurses' triage decisions. Additionally, knowledge about the time of the triage call could have resulted in bias, as the assessors may have a different threshold for assessing a decision as accurate during nighttime versus daytime. However, as we used experienced triage physicians as assessors, who assessed a call using their clinical experience, we expect this bias to have minor influence. Also, our study was underpowered to perform multivariate analyses, so we cannot ignore potential confounding concerning the associations found. Finally, we needed to exclude calls for which the level of accuracy was assessed as not appicable, as it could both reflect a correct performance

(i.e., "*this item correctly found not relevant")* or potentially cover a *poor* performance (i.e., "*available information is insufficient for assessment*").

*Interpretation and comparison of results*

Our study revealed that the triage model and time of call had an effect on accuracy of triage. In a prior study, we found that nurse-led triage had less *clinically relevant* undertriage and more *clinically relevant* overtriage than GP-led triage in randomly selected calls[23]. In this study, we found the same tendencies in a selection of potential high-risk calls. Our study could not elicit which factors of the triage models influence the difference between nurse-led and GP-led triage in high-risk calls, such as the role of CDSS, professional background, working- and organisational conditions. Nurses at the MH-1813 were obligated to use CDSS, whereas GPs did not use CDSS. CDSSs aim to ensure consistency[31,32] and to have a high degree of safety (i.e., low level of undertriage), which consequently leads to a higher level of overtriage. Furthermore, telephone triage in OOH-PC serves as a form of gatekeeping to acute healthcare in Denmark. A qualitative study described that nurses did not consider themselves as "gatekeepers", but as "service providers"[33]. Hence, perceptions of the task at hand may differ between nurses and GPs, thereby affecting overtriage, which could be seen as a service to the callers.

We found that high-risk calls during nighttime were significantly more likely to be undertriaged than calls during day or evening. This corresponds to a study by Hayward et al., finding that patients calling during low call volume (e.g., nighttime) had a higher risk of requiring secondary care within three days after the OOH-PC contact[17]. Our study cannot elicit the mechanisms behind the increased risk of undertriage during nighttime. Fatigue of the triage professional could play a role. Moreover, a stricter triage and gatekeeping function may be conducted during nighttime due to different organisational setup with less capacity of staff and consultations.

Our study may suggest that being elderly could influence the risk of undertriage in high-risk calls. This non-significant trend corresponds to prior studies, which found that increasing age was associated with unsafe triage[14,17,21]. One could hypothesise that elderly may tend to wait longer before contacting OOH-PC, which could result in calls more at risk of being urgent.

*Implications for future research and clinical practice*

Although the risk of inaccurate triage differed between nurses and GPs, knowledge of mechanisms behind this difference are lacking and need further exploration. Calling during nighttime was associated with higher risk of being undertriaged for high-risk patient calls, which may be related to a change in available resources and an urge to increase gatekeeping. Future studies should explore the effect of using a CDSS, working in different working- and organisational conditions, or having different professional background on the level of under- and overtriage. Moreover, the influence of other patient characteristics (e.g., socioeconomically factors and age) and of health complaints presented in the call are relevant themes to study in relation to the accuracy of triage. From a clinical perspective, this study suggests that triage professionals preferably should pay extra attention when making a triage decision in calls concerning abdominal pain conducted during nighttime and that extra attention may be focused when call is concerning elderly.

## CONCLUSION

This study found that high-risk calls triaged by nurses were less likely to be undertriage and more likely to be overtriage compared to calls triaged by GPs in OOH-PC. High-risk calls conducted during nighttime were significantly more likely to be undertriaged than those during day and evening.

## Contributorship statement

All authors contributed to the development of the study protocol and design. DSG collected the calls, conducted the statistical analyses, and produced the first draft of the manuscript. DSG, AFP, LH, MBC, and FF (medical doctor) contributed to the interpretation of data and critically revised the manuscript. All authors contributed with proofreading of the manuscript.

**Acknowledgements**

The authors would like to thank the 24 assessors who participated in the assessment process and the patients who gave valuable feedback during the development of the AQTT. The authors thank the MH-1813 and the GPC organisations. A special acknowledgment to statistician Claus H Vestergaard for providing valuable feedback in analyses.

**Competing interests**

None declared.

**Ethics approval**

The National Committee on Health Research Ethics in the Central Denmark Region was consulted and found that no approval was required for this study.

**Patient consent**

Consent was not obtained as approved by the Danish Patient Safety Authority (reference number: 3-3013-1274/1). The project (ID: 200) has been approved and is registered in the Record of Processing Activities at the Research Unit for General Practice in Aarhus in accordance with the provisions of the General Data Protection Regulation (GDPR).

**Transparency statement**

The lead author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as originally planned (and, if relevant, registered) have been explained.

**Funding**

This study was supported by the Danish foundation TrygFonden (104893), Primary Health Care Research Foundation of the Central Denmark Region (*Praksisforskningsfonden* - 1-15-1-72-13-09), the Committee for Quality Improvement and Continuing Medical Education in general practice in the Central Denmark Region (*Kvalitets- og Efteruddannelsesudvalget* - 1-30-72-227-15) and the Committee of Multipractice Studies in General Practice (Multipraksisudvalget - 15/1880). Researchers were independent from funders.

**Data sharing statement**

Anonymised data are available upon reasonable request.

**References**

1 Smits M, Rutten M, Keizer E, *et al.* The Development and Performance of After-Hours Primary Care in the Netherlands: A Narrative Review. *Ann Intern Med* 2017;**166**:737–42. doi:10.7326/M16-2776

2 Christensen MB, Olesen F. Out of hours service in Denmark: evaluation five years after reform. *Bmj* 1998;**316**:1502–6.

3 Anderson A, Roland M. Potential for advice from doctors to reduce the number of patients referred to emergency departments by NHS 111 call handlers: Observational study. *BMJ Open* 2015;**5**:1–4. doi:10.1136/bmjopen-2015-009444

4 Bunn F, Byrne G, Kendall S. Telephone consultation and triage: effects on health care use and patient satisfaction. *Cochrane Database of Systematic Reviews* Published Online First: 2004. doi:10.1002/14651858.CD004180.pub2

5 Salk ED, Schriger DI, Hubbell KA, *et al.* Effect of visual cues, vital signs, and protocols on triage: A prospective randomized crossover trial. *Ann Emerg Med* 1998;**32**:655–64. doi:10.1016/S0196-0644(98)70063-0

6 Morony S, Weir K, Duncan G, *et al.* Enhancing communication skills for telehealth: development and implementation of a Teach-Back intervention for a national maternal and child health helpline in

Australia. *BMC Health Serv Res* 2018;**18**:162. doi:10.1186/s12913-018-2956-6 LK - http://wx7cf7zp2h.search.serialssolutions.com?sid=EMBASE&issn=14726963&id=doi:10.1186%2Fs12913-018-2956-6&atitle=Enhancing+communication+skills+for+telehealth%3A+development+and+implementation+of+a+Teach-Back+intervention+for+a+national+maternal+and+child+health+helpline+in+Australia&stitle=BMC+Health+Serv+Res&title=BMC+health+services+research&volume=18&issue=1&spage=162&epage=&aulast=Morony&aufirst=Suzanne&auinit=S.&aufull=Morony+S.&coden=&isbn=&pages=162-&date=

7 Agha Z, Roter D, Schapira RM. An evaluation of patient-physician communication style during telemedicine consultations. *J Med Internet Res* 2009;**11**. doi:10.2196/jmir.1193

8 Ernesäter A, Winblad U, Engström M, *et al.* Malpractice claims regarding calls to Swedish telephone advice nursing: what went wrong and why? *J Telemed Telecare* 2012;**18**:379–83. doi:10.1258/jtt.2012.120416

9 Huibers L, Smits M, Renaud V, *et al.* Safety of telephone triage in out-of-hours care: A systematic review. *Scand J Prim Health Care* 2011;**29**:198–209. doi:10.3109/02813432.2011.629150

10 Morreel S, Colliers A, Remmen R, *et al.* How accurate is telephone triage in out-of-hours care? An observational trial in real patients How accurate is telephone triage in out-of-hours care? An observational trial in real patients. *Acta Clinica Belgica International Journal of Clinical and Laboratory Medicine* Published Online First: 2020. doi:10.1080/17843286.2020.1839719

11 Hansen EH, Hunskaar S. Telephone triage by nurses in primary care out-of-hours services in Norway: An evaluation study based on written case scenarios. *BMJ Qual Saf* 2011;**20**:390–6. doi:10.1136/bmjqs.2010.040824

12 Smits M, Hanssen S, Huibers L, *et al.* Telephone triage in general practices: A written case scenario study in the Netherlands. *Scand J Prim Health Care* 2016;**34**:28–36. doi:10.3109/02813432.2016.1144431

13 Montalto M, Dunt DR, Day SE, *et al.* Testing the safety of after-hours telephone triage: Patient simulations with validated scenarios. *Australasian Emergency Nursing Journal* 2010;**13**:7–16. doi:10.1016/j.aenj.2009.11.003

14 Rørtveit S, Meland E, Hunskaar S. Changes of triage by GPs during the course of prehospital emergency situations in a Norwegian rural community. *Scand J Trauma Resusc Emerg Med* 2014;**21**:89. doi:10.1186/1757-7241-21-89

15 Kristoffersen JE. Out-of-hours primary care and the patients who die. A survey of deaths after contact with a suburban primary care out-of-hours service. *Scand J Prim Health Care* 2000;**18**:139–42. doi:10.1080/028134300453322

16 Philips H, van Bergen J, Huibers L, *et al.* Agreement on urgency assessment between secretaries and general practitioners: an observational study in out-of-hours general practice service in Belgium. *Acta Clin Belg* 2015;**3286**:1–6. doi:10.1179/2295333715Y.0000000017

17 Hayward GN, Vincent C, Lasserson DS. Predicting clinical deterioration after initial assessment in out-of-hours primary care: a retrospective service evaluation. *British Journal of General Practice* 2016;:1–8. doi:10.3399/bjgp16X687961

18 Gamst-Jensen H, Lippert FK, Egerod I. Under-triage in telephone consultation is related to non-normative symptom description and interpersonal communication: a mixed methods study. *Scand J Trauma Resusc Emerg Med* 2017;**25**:1–8. doi:10.1186/s13049-017-0390-0

19 Elliott AM, McAteer A, Heaney D, *et al.* Examining the role of Scotland's telephone advice service (NHS 24) for managing health in the community: analysis of routinely collected NHS 24 data. *BMJ Open* 2015;**5**:e007293. doi:10.1136/bmjopen-2014-007293

20 North F, Odunukan O, Varkey P. The value of telephone triage for patients with appendicitis. *J Telemed Telecare* 2011;**17**:417–20. doi:10.1258/jtt.2011.110301

21 Smits M, Huibers L, Kerssemeijer B, *et al.* Patient safety in out-of-hours primary care: a review of patient records. *BMC Health Serv Res* 2010;**10**:1–9. doi:10.1186/1472-6963-10-335

22 Kempe A, Bunik M, Ellis J, *et al.* How Safe Is Triage by an After-Hours Telephone Call Center? *Pediatrics* 2006;**118**:457–63. doi:10.1542/peds.2005-3073

23 Graversen DS, Christensen MB, Pedersen AF, *et al.* Safety, efficiency and health-related quality of telephone triage conducted by general practitioners, nurses, or physicians in out-of-hours primary care: A quasi-experimental study using the Assessment of Quality in Telephone Triage (AQTT) to assess audio. *BMC Fam Pract* 2020;**21**:1–12. doi:10.1186/s12875-020-01122-z

24 Ernesäter A, Holmström IK, Engström M. Telenurses' experiences of working with computerized decision support: Supporting, inhibiting and quality improving. *J Adv Nurs* 2009;**65**:1074–83. doi:10.1111/j.1365-2648.2009.04966.x

25 Holmström IK. Decision aid software programs in telenursing: Not used as intended? Experiences of Swedish telenurses. *Nurs Health Sci* 2007;**9**:23–8. doi:10.1111/j.1442-2018.2007.00299.x

26 Holmström IK, Kaminsky E, Lindberg Y, *et al.* Registered Nurses' experiences of using a clinical decision support system for triage of emergency calls: A qualitative interview study. Published Online First: 2020. doi:10.1111/jan.14542

27 Graversen DS, Huibers L, Christensen MB, *et al.* Communication quality in telephone triage conducted by general practitioners, nurses or physicians: A quasi-experimental study using the AQTT to assess audio-recorded telephone calls to out-of-hours primary care in Denmark. *BMJ Open* 2020;**10**:1–12. doi:10.1136/bmjopen-2019-033528

28 VIVE. Regionale lægevagter og Akuttelefonen 1813. En kortlægning med fokus på organisering, aktivitet og økonomi [Regional GP Coopratives and Medical Helpline 1813. A mapping focusing on organisation, activity and economi]. 2018. doi:https://www.vive.dk/da/udgivelser/regionale-laegevagter-og-akuttelefon-1813-10796/

29 Statistics Denmark. StatBank Denmark. http://www.statbank.dk/statbank5a/default.asp?w=1920

30 Graversen DS, Pedersen AF, Carlsen AH, *et al.* Quality of out-of-hours telephone triage by general practitioners and nurses: development and testing of the AQTT - an assessment tool measuring communication, patient safety and efficiency. *Scand J Prim Health Care* 2019;**37**:18–29. doi:10.1080/02813432.2019.1568712

31 Murdoch J, Barnes R, Pooler J, *et al.* The impact of using computer decision-support software in primary care nurse-led telephone triage: interactional dilemmas and conversational consequences. *Soc Sci Med* 2015;**126**:36–47. doi:10.1016/j.socscimed.2014.12.013

32 LT W, Zwart D, Erkelens DC, *et al.* Tinkering and overruling the computer decision support system: Working strategies of telephone triage nurses who assess the urgency of callers suspected of having an acute cardiac event. *J Clin Nurs* 2019;:1175–86. doi:10.1111/jocn.15168

33 Lindberg BH, Rebnord IK, Høye S. Phone triage nurses' assessment of respiratory tract infections–the tightrope walk between gatekeeping and service providing. A qualitative study. *Scand J Prim Health Care* 2021;**39**:139–47. doi:10.1080/02813432.2021.1908715

**Figure and Table Legends**

**Box 1.** Description of the organisation and telephone triage of the two included models for out-of-hours primary care

**Figure 1.** Flowchart of inclusion and exclusion of randomly selected and high-risk calls from the GPC and MH-1813

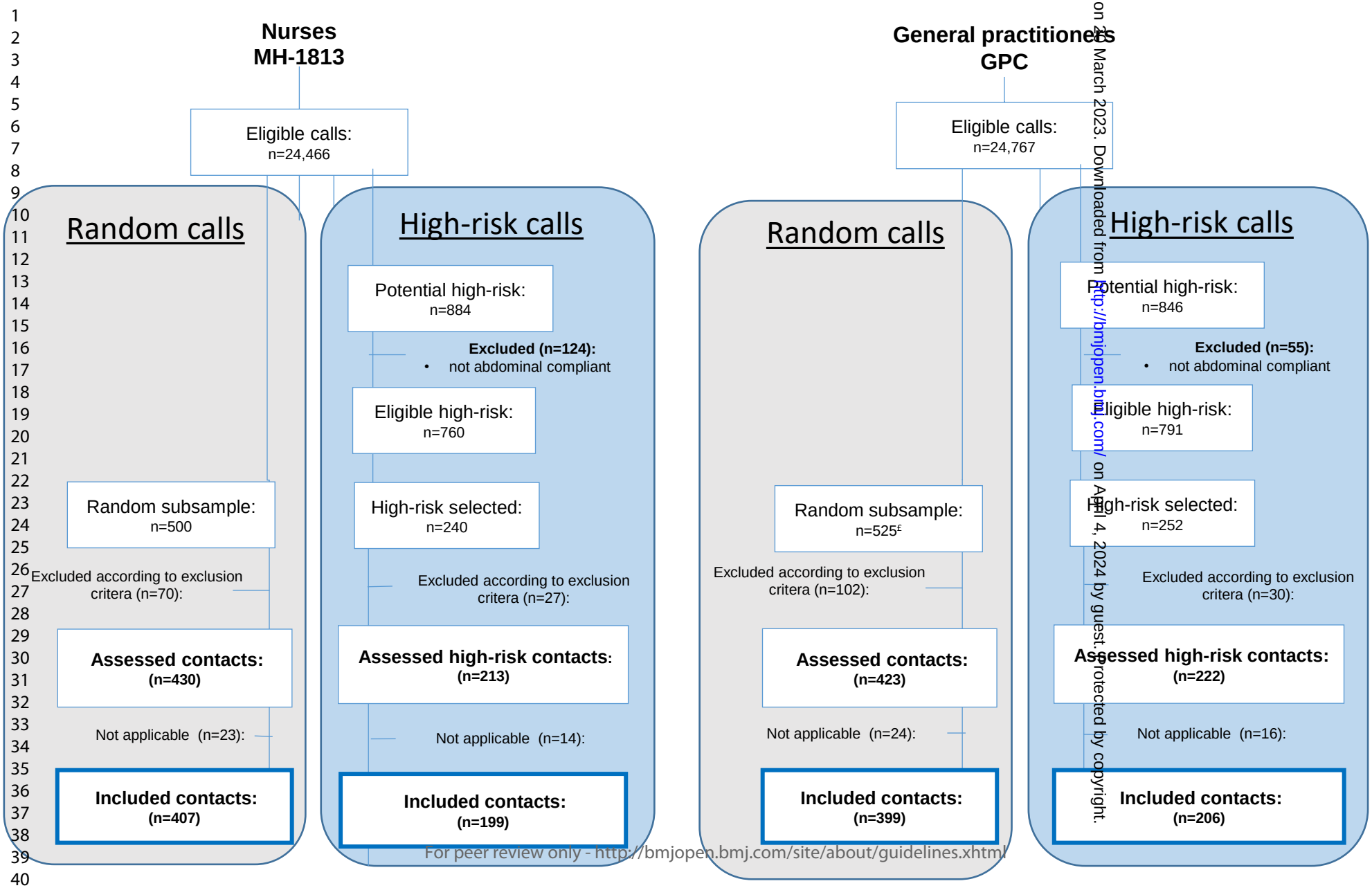
*Legends: Not applicable:* accuracy of triage could not be assessed due to insufficient information in the call or assessment of accuracy of triage was deemed not relevant.

**Table 1.** Baseline description of patient and call characteristics in randomly selected calls and high-risk calls

**Table 2.** Patient- and call characteristics associated with clinically relevant under- and overtriage in randomly selected calls[1]

**Table 3.** Patient- and call characteristics associated with clinically relevant under- and overtriage in high-risk calls

**Appendix 1.** Exclusion criteria



**Appendix 1.** Exclusion criteria

| Type of call | Definition/clarification |
|---|---|
| Frequent callers | Patients with ≥7 calls during the two-week inclusion period (the patient's medical record from the OOH service that could include important information on these patients was only available to the triage professional and not to the assessor) |
| Calls by mistake | Calls with no caller answering the triage professional |
| Daytime calls | Calls performed during daytime (the telephone triage service at MH-1813 was available during daytime) |
| Calls by other health professionals | The caller was another healthcare professional, e.g. from a nursing home |
| Administrative calls | The reason for calling was administrative, e.g. calling to get the number for the acute dentist |
| Calls regarding simple drug prescriptions | The patient called for renewal of a prescription that required little information sharing |
| Preterm termination | Calls that were ended too early, e.g. calls made by error, no sound on call, or sound interrupted in the middle of call |
| Calls from other localisation | Calls from a caller who was not in the same location as the patient, e.g. parent on the way to pick up a sick child from day care |
| Calls with poor sound quality | Calls with poor sound quality making assessment difficult |
| Language issues | Calls in which language issues challenged the triage, i.e. caller did not speak Danish or English |

| | |
|---|---|
| Not able to identify call | Calls where an exact linkage to the corresponding audio recording could not be established |

# Reporting checklist for cross sectional study.

Based on the STROBE cross sectional guidelines.

## Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the STROBE cross sectionalreporting guidelines, and cite them as:

von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies.

| | | Reporting Item | Page Number |
|---|---|---|---|
| **Title and abstract** | | | |
| Title | #1a | Indicate the study's design with a commonly used term in the title or the abstract | 1 |

| | | | |
|---|---|---|---|
| Abstract | #1b | Provide in the abstract an informative and balanced summary of what was done and what was found | 3 |
| **Introduction** | | | |
| Background / rationale | #2 | Explain the scientific background and rationale for the investigation being reported | 5 |
| Objectives | #3 | State specific objectives, including any prespecified hypotheses | 5 |
| **Methods** | | | |
| Study design | #4 | Present key elements of study design early in the paper | 5-6 |
| Setting | #5 | Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection | 5-8 |
| Eligibility criteria | #6a | Give the eligibility criteria, and the sources and methods of selection of participants. | 6-7 |
| | #7 | Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable | 7-8 |
| Data sources / measurement | #8 | For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Give information separately for for exposed and unexposed groups if applicable. | 7-8 |

| | | | |
|---|---|---|---|
| Bias | #9 | Describe any efforts to address potential sources of bias | 7-8 |
| Study size | #10 | Explain how the study size was arrived at | 6-7 |
| Quantitative variables | #11 | Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why | 8-9 |
| Statistical methods | #12a | Describe all statistical methods, including those used to control for confounding | 8-9 |
| Statistical methods | #12b | Describe any methods used to examine subgroups and interactions | 8-9 |
| Statistical methods | #12c | Explain how missing data were addressed | 9 |
| Statistical methods | #12d | If applicable, describe analytical methods taking account of sampling strategy | 9 |
| Statistical methods | #12e | Describe any sensitivity analyses | 9 |
| **Results** | | | |
| Participants | #13a | Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed. Give information separately for for exposed and unexposed groups if applicable. | 6-7 |

| | | | |
|---|---|---|---|
| Participants | #13b | Give reasons for non-participation at each stage | 6-8, appendix, flow diagram |
| Participants | #13c | Consider use of a flow diagram | 6-8, flow diagram |
| Descriptive data | #14a | Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Give information separately for exposed and unexposed groups if applicable. | 6-7,9, table 1 |
| Descriptive data | #14b | Indicate number of participants with missing data for each variable of interest | 9 |
| Outcome data | #15 | Report numbers of outcome events or summary measures. Give information separately for exposed and unexposed groups if applicable. | 9-10, Table 1,2 and 3 |
| Main results | #16a | Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included | 9-10, Table 1,2,3 |
| Main results | #16b | Report category boundaries when continuous variables were categorized | 8 |

| | | | |
|---|---|---|---|
| Main results | #16c | If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period | 9-10 |
| Other analyses | #17 | Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses | 9-10 |
| **Discussion** | | | |
| Key results | #18 | Summarise key results with reference to study objectives | 11 |
| Limitations | #19 | Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias. | 11-12 |
| Interpretation | #20 | Give a cautious overall interpretation considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence. | 12-13 |
| Generalisability | #21 | Discuss the generalisability (external validity) of the study results | 13 |
| **Other Information** | | | |
| Funding | #22 | Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based | 15 |

Notes:

- 13b: 6-8, appendix, flow diagram
- 13c: 6-8, flow diagram

- 14a: 6-7,9, table 1
- 15: 9-10, Table 1,2 and 3
- 16a: 9-10, Table 1,2,3  This checklist was completed on 19. May 2022 using https://www.goodreports.org/, a tool made by the EQUATOR Network in collaboration with Penelope.ai

BMJ Open

# Factors associated with under- and overtriage in telephone triage in Danish out-of-hours primary care – a natural quasi-experimental cross-sectional study of randomly selected and high-risk calls

| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2022-064999.R2 |
| Article Type: | Original research |
| Date Submitted by the Author: | 20-Feb-2023 |
| Complete List of Authors: | Graversen, Dennis; Aarhus University Research Unit General Practice; Aarhus University, Department of Public Health<br>Pedersen, Anette; Aarhus University Research Unit General Practice; Aarhus University, Department of clinical medicine<br>Christensen, Morten; Aarhus University Research Unit General Practice; Aarhus University, Department of Public Health<br>Folke, Fredrik; Copenhagen University Hospital, Emergency Medical Services Copenhagen; Copenhagen University Hospital, Department of Cardiology - Herlev and Gentofte<br>Huibers, L; Aarhus University Research Unit General Practice, Research Unit for General Practice |
| Primary Subject Heading: | General practice / Family practice |
| Secondary Subject Heading: | Nursing |
| Keywords: | Clinical audit < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health & safety < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, PRIMARY CARE |
| | |

*I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.*

*The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.*

*Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.*

# Factors associated with under- and overtriage in telephone triage in Danish out-of-hours primary care – a natural quasi-experimental cross-sectional study of randomly selected and high-risk calls

**Graversen DS[1,2], Pedersen AF[1,3], Christensen MB[1,2], Folke F[4,5,6], Huibers L[1]**

[1]Research Unit for General Practice, Aarhus, Denmark, [2]Department of Public Health, Aarhus University, Aarhus, Denmark, [3]Department of Clinical Medicine, Aarhus University, Aarhus, Denmark, [4]Copenhagen University Hospital - Emergency Medical Services Copenhagen, Ballerup, Denmark, [5]Department of Cardiology, Copenhagen University Hospital - Herlev and Gentofte, Denmark, [6]Department of Clinical Medicine, University of Copenhagen, Denmark

*Correspondence to*

DS Graversen (ORCID ID: 0000-0002-7784-9403), Research Unit for General Practice, Aarhus, Bartholins Allé 2, 8000 Aarhus C, Denmark & Department of Public Health, Aarhus University, Bartholins Allé 2, 8000 Aarhus C, Denmark

E-mail: d.graversen@ph.au.dk, telephone: +45 28 26 05 52

Keywords: Triage, Telephone, After-Hours Care, Primary Health Care, General Practitioners, Nurses, Safety, Efficiency, Quality of Health Care

Abstract: words 284

Word count: 2953

**Abbreviations:**

AQTT: Assessment of Quality in Telephone Triage (a validated tool for assessment of triage performance)

CDSS: Computerised decision support system

GP: General practitioners

GPC: General practitioners cooperative

MH-1813: Medical helpline 1813

OOH: Out-of-hours

OOH-PC: Out-of-hours primary care

**ABSTRACT:**

**Objectives**: We aim to explore under- and overtriage in a high-risk patient population and explore patient- and call characteristics associated with under- and overtriage in both randomly selected and in high-risk telephone calls to out-of-hours primary care (OOH-PC).

**Design:** Natural quasi-experimental cross-sectional study.

**Setting:** Two Danish OOH-PC services using different telephone triage models: a general practitioner cooperative (GPC) with GP-led triage and the medical helpline 1813 with CDSS guided nurse-led triage.

**Participants:** We included audio-recorded telephone triage calls from 2016: 806 random calls and 405 high-risk calls (defined as patients ≥ 30 years calling with abdominal pain).

**Main outcome measures:** Twenty-four experienced physicians used a validated assessment tool to assess the accuracy of triage (AQTT). We calculated the relative risk (RR) for *clinically relevant* under- and overtriage for a range of patient- and call characteristics.

**Results:** We included 806 randomly selected calls (44 *clinically relevant* undertriaged and 54 *clinically relevant* overtriaged) and 405 high-risk calls (32 undertriaged and 24 overtriaged). In high-risk calls, nurse-led triage was associated with significantly less undertriage (RR: 0.47, 95%CI: 0.23-0.97) and more overtriage (RR: 3.93, 95%CI: 1.50-10.33) compared to GP-led triage. In high-risk calls, the risk of undertriage was significantly higher for calls during nighttime (RR: 2.1, 95%CI: 1.05-4.07). Undertriage tended to be more likely for calls concerning patients ≥ 60 years compared to 30 to 59 years (11.3% vs. 6.3%) in high-risk calls. However, this result was not significant.

**Conclusion:** Nurse-led triage was associated with less undertriage and more overtriage compared to GP-led triage in high-risk calls. This study may suggest that to minimize undertriage, the triage professionals should pay extra attention when a call occurs during nighttime or concerns elderly. However, this needs confirmation in future studies.

**Strengths and limitations of this study**

- A strength of our study is the natural quasi-experimental design using real-life calls as opposed to a constructed setup.
- To our knowledge, this is the first study to explore factors associated with undertriage and overtriage in real audio-recorded calls.
- A limitation is the use of only one assessor per call, but acceptable interrater agreement was previously found.

## INTRODUCTION

In out-of-hours primary care (OOH-PC), telephone triage plays a pivotal role in managing patient flows and workload[1–3]. Telephone triage aims to ensure a safe and efficient delivery of healthcare, avoiding undertriage and minimizing overtriage[4]. However, accurate telephone triage is difficult due to the lack of visual cues of the patient, challenges in telecommunication, and time pressure[5–8].

Safety and efficiency of telephone triage in OOH-PC have been explored in a range of studies that used varying outcome measures[4,9–13]. Studies have identified a range of risk factors for potentially unsafe telephone triage in out-of-hours (OOH) care: calls concerning abdominal pain[14–20], chest pain[16,17] and shortness of breath[16,18], calls for patients with increasing age[14,17,21], and calls during nighttime[17,22]. Thus, these calls can be seen as potential high-risk calls. A few studies have explored factors associated with overtriage as a measure of inefficient telephone triage. Nurse-led triage [1,23] and triage aided by computerized decision support system (CDSS)[23–26] have been associated with overtriage. However, studies that explored factors associated with under- and overtriage used a range of study designs. None of these studies used real audio-recorded telephone calls to assess the risk of under- and overtriage and related risk factors.

Previously, we found that telephone triage was associated with significantly less undertriage and significantly more overtriage for nurse-led triage compared to GP-led triage in a sample of random calls[23] . However, it remains unclear whether this association also exists for high-risk calls. Therefore, we aim to investigate the risk of under- and overtriage in high-risk telephone calls to OOH-PC in Denmark. In addition, we aim to explore patient- and call characteristics associated with under- and overtriage in random calls and in high-risk calls to OOH-PC.

## METHODS

*Design and setting*

This paper presents secondary analyses of a natural quasi-experimental study in two regional OOH-PC services in Denmark[23,27]. In one prior paper [23], we explored safety and efficiency in randomly selected calls, using under- and overtriage. In the present paper, we included potentially high-risk calls and studied under- and overtriage of these calls. Subsequently, we explored factors associated with under- and overtriage in randomly selected calls and in high-risk patient calls.

OOH-PC is provided by the general practitioner cooperative (GPC) in the Central Denmark Region since 1992 and by the medical helpline 1813 (MH-1813) in the Capital Region of Denmark since 2014. Both services are open outside office hours (i.e., on weekdays from 4 pm to 8 am, weekends, and national holidays), offering telephone consultations, clinic consultations, and home visits. The GPC and MH-1813 use different triage models. The GPC uses GP-led telephone triage without CDSS, whereas the MH-1813 uses nurse-led telephone triage (see Box 1). At the MH-1813, the majority of incoming unselected calls (i.e., 74%) are answered by registered nurses[28]. Nurses are obliged to use a locally developed CDSS and they can redirect calls to a physician on call (i.e., 11% of all incoming unselected calls answered by a nurse)[28]. These physicians, who have various medical specialties, also conduct telephone triage. In the present paper, we excluded these calls from our analyses (i.e., triaged by physicians).

**Box 1.** Description of the organisation and telephone triage of the two included models for out-of-hours primary care

| | **GP cooperative (GPC)** | **Medical helpline 1813 (MH-1813)** |
|---|---|---|
| **Region** | **Central Denmark Region** | **Capital Region of Denmark** |
| **Population**[29] | 1.3m citizens | 1.8m citizens |

| | | |
|---|---|---|
| **Telephone calls in 2016**[28] | 701,000 | 806,000 |
| **Organiser** | GPs in the region | Regional administration |
| **Organisation and services** | ▪ Telephone consultations, home visits, and clinic consultations at the GPC<br>▪ GPs are obliged to take part in the service | ▪ Telephone consultations and home visits run by MH-1813<br>▪ Clinic consultations are located in hospital facilities and managed by EDs |
| **Remuneration of professionals** | Fee for service | Payment by the hour |
| **Triage model** | GPs or GP trainees in their final year of speciality; no CDSS available<br><br>GPs typically work 8-hour shifts and have approx. 1-4 shifts per month besides their daytime work | Nurses who are obliged to use a CDSS and have option to redirect calls to a physician<br>Physicians with different medical specialities (a minority being GPs)<br>Nurses work 8-hour shifts and are mostly fully employed at MH-1813.<br>Physicians work 8-hour shifts and are employed besides their daytime job |

*CDSS: Computerized decision support system; The triage models and organisations remain largely unchanged since 2014.*

*Definition of groups of calls*

We defined two groups of calls: 1) a group of randomly selected calls that was representative for all calls to OOH-PC and 2) a group of potentially high-risk patient calls (referred to as "high-risk"). To define the high-

risk patient calls, we conducted a systematic literature search in 2016 to identify factors associated with unsafe or inaccurate telephone triage. Seven identified studies described factors associated with a risk of unsafe telephone triage: infants[22], increasing patient age[14,17,21], calls during nighttime[17,22], abdominal pain[14–18], chest pain[16,17], and breathing difficulties[16,18]. In a consensus meeting, the authors (DSG – medical doctor, AFP - psychologist, MBC - general practitioner, LH - medical doctor) defined criteria for high-risk calls as calls concerning patients above 30 years who suffered from abdominal pain. We added the age criterion, as we aimed to include potentially dangerous conditions that present with vague, indistinctly, or greatly differing symptomatology (such as dissecting aorta aneurism, myocardial infarction, cholecystitis, pyelonephritis, acute pancreatitis, gastrointestinal ulcer, and gynaecological causes). Due to pragmatic considerations, we focused on one reason for encounter; selecting high-risk calls had to be done by manual development of an algorithm (see further).

*Selection of calls*

All calls to the GPC and MH-1813 outside office hours during the inclusion period (GPC: 23 November - 7 December 2016; MH-1813: 23 November - 8 December 2016) were potentially eligible. For calls redirected by a nurse to a physician at MH-1813, only the part conducted by the nurse was eligible and later assessed as described below. Randomly selected calls were selected from all eligible calls; these calls were the same as we studied in a prior paper [23]. The aim was to include 435 calls by GPs and 435 by nurses, based on a power calculation to detect a 5% difference in undertriage between nurse-led and GP-led undertriage in line with the aim of the main study[23]. Based on expected exclusion, we randomly selected 525 random calls from GPC and 500 calls from MH-1813.

To identify potential eligible high-risk calls, we received electronic patient journal records from the GPC and MH-1813. We defined a list of wordings and abbreviations associated with abdominal pain, which was used to search the patient records and identify all eligible calls concerning abdominal pain. DSG (medical doctor)

assessed these marked patient records to check whether inclusion for abdominal pain was correct. Calls were excluded when the triage professional in the patient record dismissed the presence of abdominal pain (i.e., triagist noted *no abdominal pain*) or when the complaint was clearly outside the thoracic, abdominal, or pelvic region. Thus, we identified 846 (GPC) and 884 (MH-1813) eligible high-risk calls for potential inclusion. A power calculation was designed to detect a significant difference between nurse-led and GP-led undertriage for high-risk calls, without detecting associations between risk factors and over-and undertriage. As this calculation revealed the need of 206 calls per triage model, we randomly selected 252 calls from the GPC and 240 calls from MH-1813. Selection of both randomly selected and high-risk calls was done, matching the overall distribution on day of week (i.e., weekend, not weekend) and time of day (i.e., day, evening, night).

Each selected call had a unique identification number that was used to identify the corresponding audio-recording. Three master students of medicine masked the audio-recordings using beep tones to cover information revealing triage profession, OOH organisation, and patient identification information. In addition, the students screened all calls for fulfilling exclusion criteria as shown in Appendix 1. DSG reviewed all calls that fulfilled exclusion criteria and those that were unclear. In case of doubt, DSG and AFP reached consensus regarding exclusion.

*Assessment of accuracy of triage*

All included calls have been assessed as described in prior papers,[23,30] using the tool "Assessment of Quality in Telephone Triage" (AQTT). The AQTT comprises 24 items assessing the health-related quality and the quality of communication[30]. AQTT has a rating manual, describing when to assign the different ratings for each item[30]. In the present study, the outcome accuracy of triage was measured by one item that used a 7-point scale to differentiate between levels of undertriage (ratings 1 to 3), optimal triage (rating 4), and overtriage (ratings 5 to 7). The AQTT showed inter-rater disagreement when using the entire

7-point scale for assessment of the accuracy of triage, but revealed satisfactory inter-rater agreement when using dichotomous scales for *clinically relevant* undertriage (ratings "1" and "2" versus ratings "3" to "7") and *clinically relevant* overtriage (ratings "6" and "7" versus ratings "1" to "5")[30].

For calls that triage nurses redirected to a physician, only the part conducted by the nurse was available for assessment. These calls could be assessed as optimal, if the decision to redirect the call was what would be expected by a nurse.

For the assessment panel, we recruited 24 physicians among triage professionals from the GPC and MH-1813 using two inclusion criteria: 1) >1 year experience and 2) active in telephone triage in OOH-PC at time of study. We randomly selected 16 GPs from the 56 interested GPs from the GPC. At MH-1813, we included all eight physicians fulfilling our inclusion criteria from the ten interested physicians. All assessors followed a two-day training course providing knowledge on telephone triage and communication, introducing the AQTT and rating manual, and assessing triage calls individually and in plenary, focusing on achieving consistency. We randomly distributed calls to all assessors, so each member of the assessment panel assessed random and high-risk calls triaged with both the GP-led model and the nurse-led model. Assessors were blinded for the type of call and triage model. Information on age and sex of the patient, day of week, and the time of each call was available for the assessor, extracted from the registration systems from the GPC and MH-1813.

*Statistical analyses*

Accuracy of triage decision was categorised into *clinically relevant* undertriage (rated "1" or "2") and *clinically relevant* overtriage (rated "6" or "7"). We conducted separate analyses for randomly selected calls and for high-risk calls. We used descriptive analyses to describe patient- and call characteristics, as well as the risk of *clinically relevant* undertriage and *clinically relevant* overtriage for both type of calls. We

explored the association of individual patient characteristics (i.e., age, patient sex) and call characteristics (i.e., weekend, time of day, triage model) with inaccurate telephone triage, by calculating the risk ratio (RR) of having *clinically relevant* undertriage (vs. no *clinically relevant* undertriage) and *clinically relevant* overtriage (vs. no *clinically relevant* overtriage), using binomial regression. 95% confidence intervals (95%CI) were calculated. All analyses were performed in Stata 14.2 (StataCorp. 2015. *Stata Statistical Software: Release 14.2*. College Station, TX: StataCorp LP).

## PUBLIC AND PATIENT INVOLVEMENT STATEMENT

We explored patients' perspective in a focus group interview concerning the development of the AQTT and incorporated received input in the rating manual of the AQTT [30].

## RESULTS

*Description of calls*

Figure 1 shows the flowchart of selection and exclusion of calls. We excluded 47 randomly selected calls and 30 high-risk calls assigned 'not applicable', as assessing accuracy of triage was not possible (e.g., insufficient information was available) or not relevant. Thus, our final study population included 806 randomly selected calls and 405 high-risk calls (Table 1). In high-risk calls, the risk of *clinically relevant* undertriage was 7.9%, whereas the risk of *clinically relevant* overtriage was 5.9%.

*(Figure 1 – flowchart- attached separate)*

**Figure 1.** Flowchart of inclusion and exclusion of randomly selected and high-risk calls from the GPC and MH-1813

*Not applicable:* accuracy of triage could not be assessed due to insufficient information in the call or assessment of accuracy of triage was deemed not relevant.

**Table 1.** Baseline description of patient and call characteristics in randomly selected calls and high-risk calls

| | **Randomly selected calls** | **High-risk calls[1]** |
|---|---|---|
| | **N=806[2]** | **N=405[2]** |
| ***Patient characteristics*** | % (N) | % (N) |
| **Age groups (in years)** | | |
| - <18 | 36.5 (294) | N.a.[1] |
| - 18-29 | 21.0 (169) | N.a.[1] |
| - 30-59 | 27.5 (222) | 67.2 (272) |
| - ≥60 | 15.0 (121) | 32.8 (133) |
| **Sex** | | |
| - Male | 40.1 (323) | 40.3 (163) |
| - Female | 59.9 (483) | 59.8 (242) |
| ***Call characteristics*** | | |
| **Weekend** | | |
| - Not weekend | 46.4 (374) | 51.4 (208) |
| - Weekend | 53.6 (432) | 48.6 (197) |
| **Time of day** | | |
| - Day or evening | 83.6 (674) | 77.5 (314) |
| - Nighttime | 16.4 (132) | 22.5 (91) |
| **Assessed accuracy of triage[3]** | | |
| *Clinically relevant* undertriage | | |

| | | |
|---|---|---|
| 1: Severe undertriage | 1.5 (12) | 3.0 (12) |
| 2: Moderate undertriage | 4.0 (32) | 4.9 (20) |
| *Satisfactory* triage | | |
| 3: Mild undertriage | 8.6 (69) | 11.6 (47) |
| 4: Optimal triage | 69.0 (556) | 66.2 (268) |
| 5: Mild overtriage | 10.3 (83) | 8.4 (34) |
| *Clinically relevant* overtriage | | |
| 6: Moderate overtriage | 4.6 (37) | 4.4 (18) |
| 7: Severe overtriage | 2.1 (17) | 1.5 (6) |
| **Triage model** | | |
| - GP-led triage | 49.5 (399) | 50.9 (206) |
| - Nurse-led triage with CDSS | 50.5 (407) | 49.1 (199) |

[1]*High-risk calls only included patients aged ≥30 years calling OOH-PC with abdominal pain.* [2]*Not applicable: Calls where the assessment of accuracy of triage was not possible or not relevant were excluded (randomly selected: n=47, high-risk; n=30).* [3]*Accuracy of triage was categorized into three categories: 1) satisfactory triage (including optimal triage ("4"), mild undertriage ("3"), and mild overtriage ('5'), 2) clinically relevant undertriage ("1" or "2"), and 3) clinically relevant overtriage ("6" or "7").*
*Definitions: Weekend = Friday 4 PM to Monday 8 AM and holidays. Day = 8 AM to 4 PM, Evening = 4 PM to midnight, Night = midnight to 8 AM.*

*Risk factors for inaccurate triage in randomly selected calls*

For randomly selected calls, we found no significant association between patient characteristics (i.e., age and sex), call characteristics (i.e., weekend and time of call) and the risk of clinically relevant under- or overtriage (Table 2).

**Table 2.** Patient- and call characteristics associated with clinically relevant under- and overtriage in randomly selected calls[1]

| | Undertriage[2] | | Overtriage[3] | |
|---|---|---|---|---|
| | N=44 | | N=54 | |
| | % (N) | RR[2] (95%CI) | % (N) | RR[2] (95%CI) |
| ***Patient characteristics*** | | | | |
| **Age groups (in years)** | | | | |
| - <18 (n=294) | 3.7 (11) | 0.57 (0.23-1.37) | 7.5 (22) | 1.51 (0.63-3.63) |
| - 18-29 (n=169) | 4.1 (7) | 0.62 (0.23-1.68) | 7.7 (13) | 1.55 (0.61-3.97) |
| - 30-59 (n=222) | 8.1 (18) | 1.23 (0.55-2.74) | 5.9 (13) | 1.18 (0.46-3.03) |
| - ≥ 60 (n=121) | 6.6 (8) | 1 | 5.0 (6) | 1 |
| **Sex** | | | | |
| - Male (n=323) | 5.3 (17) | 1 | 4.6 (15) | 1 |
| - Female (n=483) | 5.6 (27) | 1.06 (0.59-1.92) | 8.1 (39) | 1.74 (0.97-3.10) |
| ***Call characteristics*** | | | | |
| **Weekend** | | | | |
| - Not weekend (n=374) | 5.4 (20) | 1 | 6.7 (25) | 1 |

| | | | | |
|---|---|---|---|---|
| - Weekend (n=432) | 5.6 (24) | 1.04 (0.58-1.85) | 6.7 (29) | 1.00 (0.60-1.68) |
| **Time of day** | | | | |
| - Day or evening (n=674) | 5.3 (36) | 1 | 7.0 (47) | 1 |
| - Nighttime (n=132) | 6.1 (8) | 1.13 (0.54-2.39) | 5.3 (7) | 0.76 (0.35-1.65) |

*[1]We excluded 47 calls, as assessing accuracy of triage was not possible or not relevant; [2]Undertriage: triage decision assessed as clinically relevant undertriage (rated "1" or "2"); [3]Overtriage: clinically relevant overtriage (rated "6" or "7") ; *Significant difference in RR, using binomial regression analyses; p<0.05: Definitions: Weekend = Friday 4 PM to Monday 8 AM and holidays. Day = 8 AM to 4 PM, Evening = 4 PM to midnight, Night = midnight to 8 AM.*



*Risk factors for inaccurate triage in high-risk calls*

Nurse-led triage was associated with significantly less *clinically relevant* undertriage than GP-led triage (RR=0.47, 95%CI: 0.23-0.97) in high-risk calls (Table 3). Nurse-led triage had significantly more *clinically relevant* overtriage (RR=3.93, 95%CI: 1.50-10.53) compared to GP-led triage. High-risk calls conducted during nighttime had a significantly higher risk of being *clinically relevant* undertriaged (13.2%) compared to calls conducted during day or evening (6.4%) (RR=2.1, 95%CI: 1.05-4.07). We found no significant association for the other patient- and call characteristics. However, a trend was seen for patient age, as the

risk of *clinically relevant* undertriage was lower in patients aged 30 to 59 years (6.3%) in comparison with elderly patients ≥60 years (11.3%) (RR=0.55, 95%CI: 0.29-1.07).

**Table 3.** Patient- and call characteristics associated with clinically relevant under- and overtriage in high-risk calls

| | **Undertriage[1]** | | **Overtriage[2]** | |
|---|---|---|---|---|
| | **N=32** | | **N=24** | |
| | % (N) | RR (95%CI) | % (N) | RR (95%CI) |
| ***Patient characteristics*** | | | | |
| **Age groups (in years)** | | | | |
| - 30-59 (n=272) | 6.3 (17) | 0.55 (0.29-1.07) | 7.4 (20) | 2.44 (0.85-7.01) |
| - >60 (n=133) | 11.3 (15) | 1 | 3.0 (4) | 1 |
| **Sex** | | | | |
| - Male (n=163) | 8.6 (14) | 1 | 6.8 (11) | 1 |
| - Female (n=242) | 7.4 (18) | 0.87 (0.44-1.69) | 5.4 (13) | 0.80 (0.40-1.73) |
| ***Call characteristics*:** | | | | |
| **Weekend** | | | | |
| - Not weekend (n=208) | 6.7 (14) | 1 | 6.3 (13) | 1 |
| - Weekend (n=197) | 9.1 (18) | 1.36 (0.69-2.65) | 5.6 (11) | 0.89 (0.41-1.95) |
| **Time of day** | | | | |

| | | | | |
|---|---|---|---|---|
| - Day or evening (n=314)<br>- Nighttime (n=91) | 6.4 (20)<br>13.2 (12) | 1<br>2.1 (1.05-4.07)* | 6.1 (19)<br>5.5 (5) | 1<br>0.91 (0.35-2.36) |
| **Triage model**<br>- GP-led triage (n=206)<br>- Nurse-led triage with CDSS (n=199) | 10.7 (22)<br>5.0 (10) | 1<br>0.47 (0.23-0.97)* | 2.4 (5)<br>10.0 (19) | 1<br>3.93 (1.50-10.33)* |

*We excluded 30 calls as assessing accuracy of triage was not possible or not relevant; [1]Undertriage: triage decision assessed as clinically relevant undertriage (rated "1" or "2"): [2]Overtriage: clinically relevant overtriage (rated "6" or "7"); *Significant difference in RR, using binomial regression analyses; p<0.05: Definitions: Weekend = Friday 4 PM to Monday 8 AM and holidays. Day = 8 AM to 4 PM, Evening = 4 PM to midnight, Night = midnight to 8 AM.*

## DISCUSSION

### *Principal findings*

In high-risk calls, nurse-led triage was associated with significantly lower risk of *clinically relevant* undertriage and significant higher risk of *clinically relevant* overtriage compared to GP-led triage. For high-risk calls, the risk of *clinically relevant* undertriage was significantly higher if the call was conducted during nighttime compared to day and evening. For randomly selected calls, we found no significant association between defined risk factors and the risk of *clinically relevant* under- or overtriage

### *Interpretation and comparison of results*

Our study revealed that the triage model and time of call had an effect on accuracy of triage. In a prior study, we found that nurse-led triage had less *clinically relevant* undertriage and more *clinically relevant* overtriage than GP-led triage in randomly selected calls[23]. In this study, we found the same tendencies in

a selection of potential high-risk calls. Our study could not elicit which factors of the triage models influence the difference between nurse-led and GP-led triage in high-risk calls, such as the role of CDSS, professional background, working- and organisational conditions. Nurses at the MH-1813 were obligated to use CDSS, whereas GPs did not use CDSS. CDSSs aim to ensure consistency[31,32] and to have a high degree of safety (i.e., low level of undertriage), which consequently leads to a higher level of overtriage. Furthermore, telephone triage in OOH-PC serves as a form of gatekeeping to acute healthcare in Denmark. A qualitative study described that nurses did not consider themselves as "gatekeepers", but as "service providers"[33]. Hence, perceptions of the task at hand may differ between nurses and GPs, thereby affecting overtriage, which could be seen as a service to the callers.

We found that high-risk calls during nighttime were significantly more likely to be undertriaged than calls during day or evening. This corresponds to a study by Hayward et al., finding that patients calling during low call volume (e.g., nighttime) had a higher risk of requiring secondary care within three days after the OOH-PC contact[17]. Our study cannot elicit the mechanisms behind the increased risk of undertriage during nighttime. Fatigue of the triage professional could play a role. Moreover, a stricter triage and gatekeeping function may be conducted during nighttime due to different organisational setup with less capacity of staff and consultations.

Our study may suggest that being elderly could influence the risk of undertriage in high-risk calls. This non-significant trend corresponds to prior studies, which found that increasing age was associated with unsafe triage[14,17,21]. One could hypothesise that elderly may tend to wait longer before contacting OOH-PC, which could result in calls more at risk of being urgent.

*Implications for future research and clinical practice*

Although the risk of inaccurate triage differed between nurses and GPs, knowledge of mechanisms behind this difference are lacking and need further exploration. Calling during nighttime was associated with higher risk of being undertriaged for high-risk patient calls, which may be related to a change in available resources and an urge to increase gatekeeping. Future studies should explore the effect of using a CDSS, working in different working- and organisational conditions, or having different professional background on the level of under- and overtriage. Moreover, the influence of other patient characteristics (e.g., socioeconomically factors and age) and of health complaints presented in the call are relevant themes to study in relation to the accuracy of triage. From a clinical perspective, this study suggests that triage professionals preferably should pay extra attention when making a triage decision in calls concerning abdominal pain conducted during nighttime and that extra attention may be focused when call is concerning elderly.

*Strengths and limitations of the study*

A major strength was the quasi-experimental design using real-life calls as opposed to the constructed setup used in previous studies [18–20, 31, 32]. An additional strength was the meticulous assessment process using the validated AQTT tool combined with a comprehensive rating manual [30].

Our study had the following limitations. Due to the thorough assessment process, each call was assessed by one assessor. Consequently, bias due to misclassification cannot be rejected. However, we took several precautions to ensure consistency of assessments. The assessors followed a comprehensive training course and assessments followed the meticulously developed and validated AQTT[30]. Furthermore, we attempted to mask the audio-recordings for information about organisation and triage model. Also, assessors were not aware of the design with both randomly selected and high-risk calls. Moreover, we dichotomised accuracy of triage into *clinically relevant* undertriage and overtriage, which had a satisfactory inter-rater agreement of the AQTT [30]. We carefully decided upon our definition of high-risk calls (i.e., calls for adults >30 years with abdominal pain). The cut-off point for age was reached through meticulous

discussions among the authors, but could be too low, thereby including a larger group of low-risk calls. Another limitation was our small sample size. In line with the main study, our power calculation was made to identify a significant difference between nurse-led and GP-led under- and overtriage. As only a selection of these calls was assessed as *clinically relevant* under- or overtriage, the present study lacked power to identify risk factors for under- and overtriage. Therefore, we explored clinically relevant associations of patient- and call characteristics with inaccurate triage. Furthermore, we chose physicians as assessors of the accuracy of triage, as they were most frequently used in other studies. The decision to include only physicians in the assessment panel may have induced similar-to-me cognitive bias when assessing nurse-led triage, leading to underassessment of the accuracy of nurses' triage decisions. Additionally, knowledge about the time of the triage call could have resulted in bias, as the assessors may have a different threshold for assessing a decision as accurate during nighttime versus daytime. However, as we used experienced triage physicians as assessors, who assessed a call using their clinical experience, we expect this bias to have minor influence. Also, our study was underpowered to perform multivariate analyses, so we cannot ignore potential confounding concerning the associations found. Finally, we needed to exclude calls for which the level of accuracy was assessed as not appicable, as it could both reflect a correct performance (i.e., "*this item correctly found not relevant")* or potentially cover a *poor* performance (i.e., "*available information is insufficient for assessment*").

## CONCLUSION

This study found that high-risk calls triaged by nurses were less likely to be undertriage and more likely to be overtriage compared to calls triaged by GPs in OOH-PC. High-risk calls conducted during nighttime were significantly more likely to be undertriaged than those during day and evening.

## Contributorship statement

All authors contributed to the development of the study protocol and design. DSG collected the calls, conducted the statistical analyses, and produced the first draft of the manuscript. DSG, AFP, LH, MBC, and FF (medical doctor) contributed to the interpretation of data and critically revised the manuscript. All authors contributed with proofreading of the manuscript.

**Acknowledgements**

The authors would like to thank the 24 assessors who participated in the assessment process and the patients who gave valuable feedback during the development of the AQTT. The authors thank the MH-1813 and the GPC organisations. A special acknowledgment to statistician Claus H Vestergaard for providing valuable feedback in analyses.

**Competing interests**

None declared.

**Ethics approval**

The National Committee on Health Research Ethics in the Central Denmark Region was consulted and found that no approval was required for this study.

**Patient consent**

Consent was not obtained as approved by the Danish Patient Safety Authority (reference number: 3-3013-1274/1). The project (ID: 200) has been approved and is registered in the Record of Processing Activities at the Research Unit for General Practice in Aarhus in accordance with the provisions of the General Data Protection Regulation (GDPR).

**Transparency statement**

The lead author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as originally planned (and, if relevant, registered) have been explained.

**Funding**

This study was supported by the Danish foundation TrygFonden (104893), Primary Health Care Research Foundation of the Central Denmark Region (*Praksisforskningsfonden* - 1-15-1-72-13-09), the Committee for Quality Improvement and Continuing Medical Education in general practice in the Central Denmark Region (*Kvalitets- og Efteruddannelsesudvalget* - 1-30-72-227-15) and the Committee of Multipractice Studies in General Practice (Multipraksisudvalget - 15/1880). Researchers were independent from funders.

**Data sharing statement**

Anonymised data are available upon reasonable request.

**References**

1 Smits M, Rutten M, Keizer E, *et al.* The Development and Performance of After-Hours Primary Care in the Netherlands: A Narrative Review. *Ann Intern Med* 2017;**166**:737–42. doi:10.7326/M16-2776

2 Christensen MB, Olesen F. Out of hours service in Denmark: evaluation five years after reform. *Bmj* 1998;**316**:1502–6.

3 Anderson A, Roland M. Potential for advice from doctors to reduce the number of patients referred to emergency departments by NHS 111 call handlers: Observational study. *BMJ Open* 2015;**5**:1–4. doi:10.1136/bmjopen-2015-009444

4 Bunn F, Byrne G, Kendall S. Telephone consultation and triage: effects on health care use and patient satisfaction. *Cochrane Database of Systematic Reviews* Published Online First: 2004. doi:10.1002/14651858.CD004180.pub2

5 Salk ED, Schriger DI, Hubbell KA, *et al.* Effect of visual cues, vital signs, and protocols on triage: A prospective randomized crossover trial. *Ann Emerg Med* 1998;**32**:655–64. doi:10.1016/S0196-0644(98)70063-0

6 Morony S, Weir K, Duncan G, *et al.* Enhancing communication skills for telehealth: development and implementation of a Teach-Back intervention for a national maternal and child health helpline in

Australia. *BMC Health Serv Res* 2018;**18**:162. doi:10.1186/s12913-018-2956-6 LK - http://wx7cf7zp2h.search.serialssolutions.com?sid=EMBASE&issn=14726963&id=doi:10.1186%2Fs12913-018-2956-6&atitle=Enhancing+communication+skills+for+telehealth%3A+development+and+implementation+of+a+Teach-Back+intervention+for+a+national+maternal+and+child+health+helpline+in+Australia&stitle=BMC+Health+Serv+Res&title=BMC+health+services+research&volume=18&issue=1&spage=162&epage=&aulast=Morony&aufirst=Suzanne&auinit=S.&aufull=Morony+S.&coden=&isbn=&pages=162-&date=

7 Agha Z, Roter D, Schapira RM. An evaluation of patient-physician communication style during telemedicine consultations. *J Med Internet Res* 2009;**11**. doi:10.2196/jmir.1193

8 Ernesäter A, Winblad U, Engström M, *et al.* Malpractice claims regarding calls to Swedish telephone advice nursing: what went wrong and why? *J Telemed Telecare* 2012;**18**:379–83. doi:10.1258/jtt.2012.120416

9 Huibers L, Smits M, Renaud V, *et al.* Safety of telephone triage in out-of-hours care: A systematic review. *Scand J Prim Health Care* 2011;**29**:198–209. doi:10.3109/02813432.2011.629150

10 Morreel S, Colliers A, Remmen R, *et al.* How accurate is telephone triage in out-of-hours care? An observational trial in real patients How accurate is telephone triage in out-of-hours care? An observational trial in real patients. *Acta Clinica Belgica International Journal of Clinical and Laboratory Medicine* Published Online First: 2020. doi:10.1080/17843286.2020.1839719

11 Hansen EH, Hunskaar S. Telephone triage by nurses in primary care out-of-hours services in Norway: An evaluation study based on written case scenarios. *BMJ Qual Saf* 2011;**20**:390–6. doi:10.1136/bmjqs.2010.040824

12 Smits M, Hanssen S, Huibers L, *et al.* Telephone triage in general practices: A written case scenario study in the Netherlands. *Scand J Prim Health Care* 2016;**34**:28–36. doi:10.3109/02813432.2016.1144431

13 Montalto M, Dunt DR, Day SE, *et al.* Testing the safety of after-hours telephone triage: Patient simulations with validated scenarios. *Australasian Emergency Nursing Journal* 2010;**13**:7–16. doi:10.1016/j.aenj.2009.11.003

14 Rørtveit S, Meland E, Hunskaar S. Changes of triage by GPs during the course of prehospital emergency situations in a Norwegian rural community. *Scand J Trauma Resusc Emerg Med* 2014;**21**:89. doi:10.1186/1757-7241-21-89

15 Kristoffersen JE. Out-of-hours primary care and the patients who die. A survey of deaths after contact with a suburban primary care out-of-hours service. *Scand J Prim Health Care* 2000;**18**:139–42. doi:10.1080/028134300453322

16 Philips H, van Bergen J, Huibers L, *et al.* Agreement on urgency assessment between secretaries and general practitioners: an observational study in out-of-hours general practice service in Belgium. *Acta Clin Belg* 2015;**3286**:1–6. doi:10.1179/2295333715Y.0000000017

17 Hayward GN, Vincent C, Lasserson DS. Predicting clinical deterioration after initial assessment in out-of-hours primary care: a retrospective service evaluation. *British Journal of General Practice* 2016;:1–8. doi:10.3399/bjgp16X687961

18 Gamst-Jensen H, Lippert FK, Egerod I. Under-triage in telephone consultation is related to non-normative symptom description and interpersonal communication: a mixed methods study. *Scand J Trauma Resusc Emerg Med* 2017;**25**:1–8. doi:10.1186/s13049-017-0390-0

19 Elliott AM, McAteer A, Heaney D, *et al.* Examining the role of Scotland's telephone advice service (NHS 24) for managing health in the community: analysis of routinely collected NHS 24 data. *BMJ Open* 2015;**5**:e007293. doi:10.1136/bmjopen-2014-007293

20 North F, Odunukan O, Varkey P. The value of telephone triage for patients with appendicitis. *J Telemed Telecare* 2011;**17**:417–20. doi:10.1258/jtt.2011.110301

21 Smits M, Huibers L, Kerssemeijer B, *et al.* Patient safety in out-of-hours primary care: a review of patient records. *BMC Health Serv Res* 2010;**10**:1–9. doi:10.1186/1472-6963-10-335

22 Kempe A, Bunik M, Ellis J, *et al.* How Safe Is Triage by an After-Hours Telephone Call Center? *Pediatrics* 2006;**118**:457–63. doi:10.1542/peds.2005-3073

23 Graversen DS, Christensen MB, Pedersen AF, *et al.* Safety, efficiency and health-related quality of telephone triage conducted by general practitioners, nurses, or physicians in out-of-hours primary care: A quasi-experimental study using the Assessment of Quality in Telephone Triage (AQTT) to assess audio. *BMC Fam Pract* 2020;**21**:1–12. doi:10.1186/s12875-020-01122-z

24 Ernesäter A, Holmström IK, Engström M. Telenurses' experiences of working with computerized decision support: Supporting, inhibiting and quality improving. *J Adv Nurs* 2009;**65**:1074–83. doi:10.1111/j.1365-2648.2009.04966.x

25 Holmström IK. Decision aid software programs in telenursing: Not used as intended? Experiences of Swedish telenurses. *Nurs Health Sci* 2007;**9**:23–8. doi:10.1111/j.1442-2018.2007.00299.x

26 Holmström IK, Kaminsky E, Lindberg Y, *et al.* Registered Nurses' experiences of using a clinical decision support system for triage of emergency calls: A qualitative interview study. Published Online First: 2020. doi:10.1111/jan.14542

27 Graversen DS, Huibers L, Christensen MB, *et al.* Communication quality in telephone triage conducted by general practitioners, nurses or physicians: A quasi-experimental study using the AQTT to assess audio-recorded telephone calls to out-of-hours primary care in Denmark. *BMJ Open* 2020;**10**:1–12. doi:10.1136/bmjopen-2019-033528

28 VIVE. Regionale lægevagter og Akuttelefonen 1813. En kortlægning med fokus på organisering, aktivitet og økonomi [Regional GP Coopratives and Medical Helpline 1813. A mapping focusing on organisation, activity and economi]. 2018. doi:https://www.vive.dk/da/udgivelser/regionale-laegevagter-og-akuttelefon-1813-10796/

29 Statistics Denmark. StatBank Denmark. http://www.statbank.dk/statbank5a/default.asp?w=1920

30 Graversen DS, Pedersen AF, Carlsen AH, *et al.* Quality of out-of-hours telephone triage by general practitioners and nurses: development and testing of the AQTT - an assessment tool measuring communication, patient safety and efficiency. *Scand J Prim Health Care* 2019;**37**:18–29. doi:10.1080/02813432.2019.1568712

31 Murdoch J, Barnes R, Pooler J, *et al.* The impact of using computer decision-support software in primary care nurse-led telephone triage: interactional dilemmas and conversational consequences. *Soc Sci Med* 2015;**126**:36–47. doi:10.1016/j.socscimed.2014.12.013

32 LT W, Zwart D, Erkelens DC, *et al.* Tinkering and overruling the computer decision support system: Working strategies of telephone triage nurses who assess the urgency of callers suspected of having an acute cardiac event. *J Clin Nurs* 2019;:1175–86. doi:10.1111/jocn.15168

33 Lindberg BH, Rebnord IK, Høye S. Phone triage nurses' assessment of respiratory tract infections–the tightrope walk between gatekeeping and service providing. A qualitative study. *Scand J Prim Health Care* 2021;**39**:139–47. doi:10.1080/02813432.2021.1908715

**Figure and Table Legends**

**Box 1.** Description of the organisation and telephone triage of the two included models for out-of-hours primary care

**Figure 1.** Flowchart of inclusion and exclusion of randomly selected and high-risk calls from the GPC and MH-1813

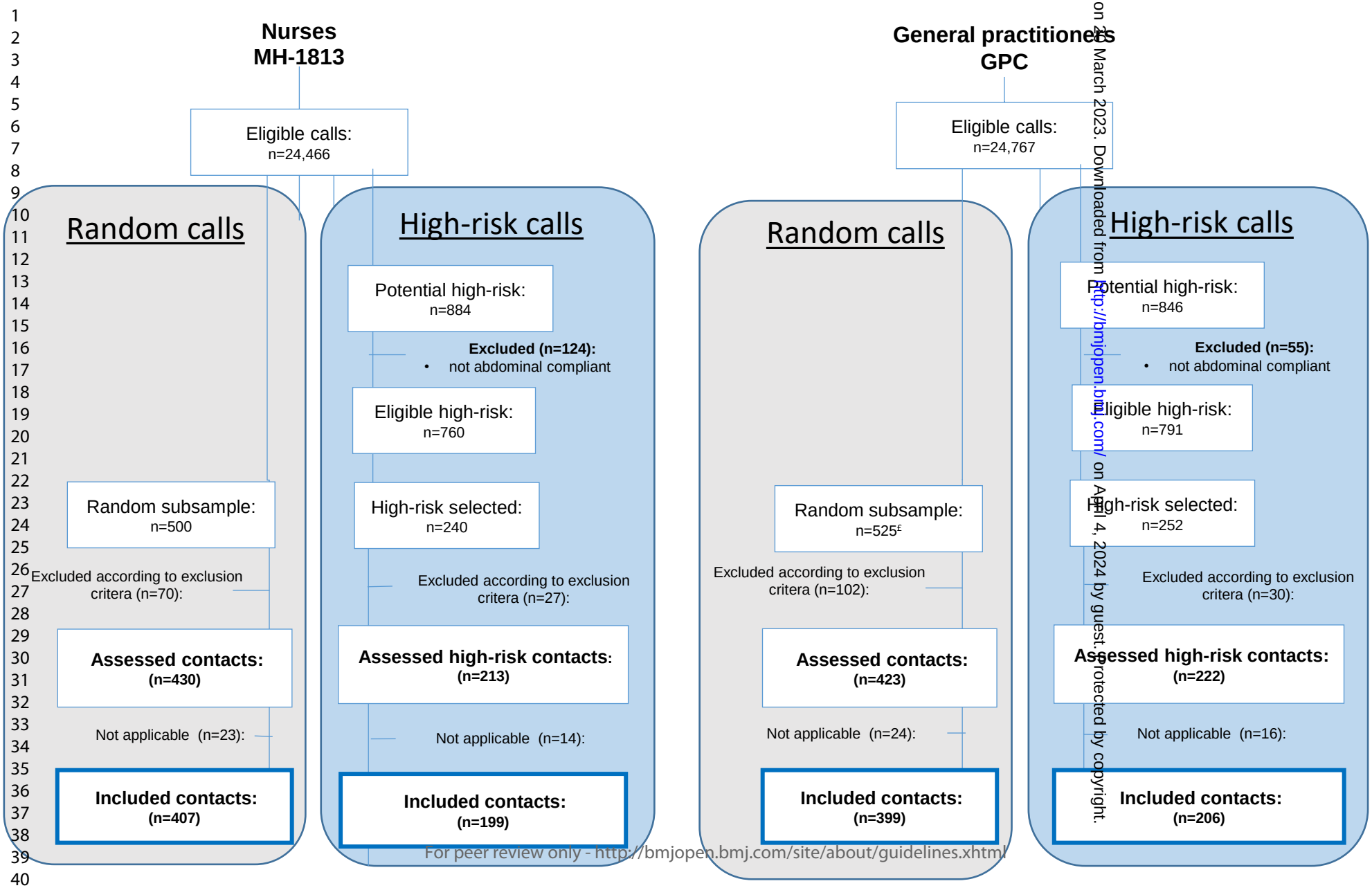
*Legends: Not applicable:* accuracy of triage could not be assessed due to insufficient information in the call or assessment of accuracy of triage was deemed not relevant.

**Table 1.** Baseline description of patient and call characteristics in randomly selected calls and high-risk calls

**Table 2.** Patient- and call characteristics associated with clinically relevant under- and overtriage in randomly selected calls[1]

**Table 3.** Patient- and call characteristics associated with clinically relevant under- and overtriage in high-risk calls

**Appendix 1.** Exclusion criteria



**Appendix 1.** Exclusion criteria

| Type of call | Definition/clarification |
|---|---|
| Frequent callers | Patients with ≥7 calls during the two-week inclusion period (the patient's medical record from the OOH service that could include important information on these patients was only available to the triage professional and not to the assessor) |
| Calls by mistake | Calls with no caller answering the triage professional |
| Daytime calls | Calls performed during daytime (the telephone triage service at MH-1813 was available during daytime) |
| Calls by other health professionals | The caller was another healthcare professional, e.g. from a nursing home |
| Administrative calls | The reason for calling was administrative, e.g. calling to get the number for the acute dentist |
| Calls regarding simple drug prescriptions | The patient called for renewal of a prescription that required little information sharing |
| Preterm termination | Calls that were ended too early, e.g. calls made by error, no sound on call, or sound interrupted in the middle of call |
| Calls from other localisation | Calls from a caller who was not in the same location as the patient, e.g. parent on the way to pick up a sick child from day care |
| Calls with poor sound quality | Calls with poor sound quality making assessment difficult |
| Language issues | Calls in which language issues challenged the triage, i.e. caller did not speak Danish or English |

| | |
|---|---|
| Not able to identify call | Calls where an exact linkage to the corresponding audio recording could not be established |

# Reporting checklist for cross sectional study.

Based on the STROBE cross sectional guidelines.

## Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the STROBE cross sectionalreporting guidelines, and cite them as:

von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies.

| | | Reporting Item | Page Number |
|---|---|---|---|
| **Title and abstract** | | | |
| Title | #1a | Indicate the study's design with a commonly used term in the title or the abstract | 1 |

| | | | |
|---|---|---|---|
| Abstract | #1b | Provide in the abstract an informative and balanced summary of what was done and what was found | 3 |
| **Introduction** | | | |
| Background / rationale | #2 | Explain the scientific background and rationale for the investigation being reported | 5 |
| Objectives | #3 | State specific objectives, including any prespecified hypotheses | 5 |
| **Methods** | | | |
| Study design | #4 | Present key elements of study design early in the paper | 5-6 |
| Setting | #5 | Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection | 5-8 |
| Eligibility criteria | #6a | Give the eligibility criteria, and the sources and methods of selection of participants. | 6-7 |
| | #7 | Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable | 7-8 |
| Data sources / measurement | #8 | For each variable of interest give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Give information separately for for exposed and unexposed groups if applicable. | 7-8 |

| | | | |
|---|---|---|---|
| Bias | #9 | Describe any efforts to address potential sources of bias | 7-8 |
| Study size | #10 | Explain how the study size was arrived at | 6-7 |
| Quantitative variables | #11 | Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why | 8-9 |
| Statistical methods | #12a | Describe all statistical methods, including those used to control for confounding | 8-9 |
| Statistical methods | #12b | Describe any methods used to examine subgroups and interactions | 8-9 |
| Statistical methods | #12c | Explain how missing data were addressed | 9 |
| Statistical methods | #12d | If applicable, describe analytical methods taking account of sampling strategy | 9 |
| Statistical methods | #12e | Describe any sensitivity analyses | 9 |
| **Results** | | | |
| Participants | #13a | Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed. Give information separately for for exposed and unexposed groups if applicable. | 6-7 |

| | | | |
|---|---|---|---|
| Participants | #13b | Give reasons for non-participation at each stage | 6-8, appendix, flow diagram |
| Participants | #13c | Consider use of a flow diagram | 6-8, flow diagram |
| Descriptive data | #14a | Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Give information separately for exposed and unexposed groups if applicable. | 6-7,9, table 1 |
| Descriptive data | #14b | Indicate number of participants with missing data for each variable of interest | 9 |
| Outcome data | #15 | Report numbers of outcome events or summary measures. Give information separately for exposed and unexposed groups if applicable. | 9-10, Table 1,2 and 3 |
| Main results | #16a | Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included | 9-10, Table 1,2,3 |
| Main results | #16b | Report category boundaries when continuous variables were categorized | 8 |

| | | | |
|---|---|---|---|
| Main results | #16c | If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period | 9-10 |
| Other analyses | #17 | Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses | 9-10 |
| **Discussion** | | | |
| Key results | #18 | Summarise key results with reference to study objectives | 11 |
| Limitations | #19 | Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias. | 11-12 |
| Interpretation | #20 | Give a cautious overall interpretation considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence. | 12-13 |
| Generalisability | #21 | Discuss the generalisability (external validity) of the study results | 13 |
| **Other Information** | | | |
| Funding | #22 | Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based | 15 |

Notes:

- 13b: 6-8, appendix, flow diagram
- 13c: 6-8, flow diagram

- 14a: 6-7,9, table 1
- 15: 9-10, Table 1,2 and 3
- 16a: 9-10, Table 1,2,3  This checklist was completed on 19. May 2022 using https://www.goodreports.org/, a tool made by the EQUATOR Network in collaboration with Penelope.ai