## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (http://bmjopen.bmj.com/site/about/resources/checklist.pdf) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

#### ARTICLE DETAILS

TITLE (PROVISIONAL)	Relation between illness representation and self-reported degree-of-
	worry in patients calling out-of-hours services: a mixed methods
	study in Copenhagen, Denmark
AUTHORS	Thilsted, Sita; Egerod, Ingrid; Lippert, Freddy; Gamst-Jensen, Hejdi

#### **VERSION 1 – REVIEW**

REVIEWER	Barbara Farquharson	
	Edinburgh Napier University, Scotland	
REVIEW RETURNED	15-Nov-2017	

· · · · · · · · · · · · · · · · · · ·	
GENERAL COMMENTS	Thank you for asking me to review this interesting paper which describes a neat study and which builds on work I published in 2011. In general, I think the data are of potential interest to the readership of BMJ Open and have the potential to be relevant to clinical practice of telephone triage. However, there are a number of aspects of the manuscript that require additional attention to ensure the reader has all the information required to evaluate the findings: 1. The main finding reported is a relationship between illness representations and self-evaluation of urgency (also referred to as degree of worry). However, the theoretical basis of this construct is not described, is measured by a single question and there is no mention of validity and reliability - is there a validated measure? If not, what data can you offer to support the validity and reliability of this measure? The reference for the intensity descriptor relates to pain - what evidence is there that it is applicable to DOW? 2. Statements are made about the relationship between constructs (e.g. strong identity and low DOW) and indeed these are the main findings, however they are not supported by any statistics which allow the reader to evaluate the strength and significance of the associations (looking at the graphs in Fig 3 I suspect some would not be significant) - please present the statistical data to support your main findings. 3. It is stated that ethical approval was not require dwhich surprised me - in the UK, a study of this nature would require ethical approval. In any case, details of the steps taken to protect participants/ensure rigor should be described in the manuscript. 4. Additional data about the participants would also be helpful to the reader - what age, gender, ethnicity etc, what was the nature of the calls, were they representative? How were the dates for participation chosen? 5. For the qualitative analysis - more detail of how coding units were identified would help replicability. More detail as to how blinding to

DOW was ensured would provide more reassurance re rigour. A limitation identified relates to a single coder having identified the components of CSM - this would be easily overcome by having an additional coder perform a reliability check - providing additional detail of this would increase confidence in the results presented. 6. The results reported relating to timeline describe the duration of symptoms prior to the call - this is of interest but the definition of this construct proposed by Leventhal relates to the perceived duration of the symptoms/condition - can you present additional data which reflects callers perceptions of timeline as well as the duration so far? The quantitative findings it is reported that callers with a low DOW were more likely to mention a cause for their illness etc. However, as it is not differentiated within the data whether the illness representation was presented spontaneously or prompted by a callhandler, it is not also possible that the relationship could be due to call-handlers being more/less likely to ask prompting questions depending on DOW? Further consideration of this would be helpful.
A strobe checklist was not available to me - was it submitted?
Minor point but might it be preferable to phrase things in a non- gendered way (e.g. 'their illness representations' rather than 'her illness representations')

REVIEWER	Ellen Keizer
	Institute of Primary Care, University Hospital Zürich, Zürich,
	Switzerland
REVIEW RETURNED	11-Jan-2018
GENERAL COMMENTS	Interrelation between illness representation and self-reported degree-of-worry in patients calling out-of-hours services: a mixed methods study. Thilsted et al. BMJ Open It is an interesting topic and generally a clearly written paper. However, the methods of this study are weak in my opinion. I don't think the method fits to the research question; if you want to find an interrelation more power is needed. I consider this study more as a pilot study. Another limitation is that only one researcher conducted the qualitative analyses. Although I think it is good to ask patients how worried they are when they contact OOH services, I don't think you can prove with this research that it will (potentially) increase patient safety, you need more evidence for that statement. I have made some specific comments below. Introduction 1. I think the rationale for this study should be elaborated further in the introduction. Is there evidence that there is a relationship between degree of worry and urgency from a medical perspective? Is worry a good indicator for assessing urgency, and thus patient safety? I think this statement need some evidence. You should refer to studies which have found a relationship between worry and
	<ul> <li>the qualitative analyses.</li> <li>Although I think it is good to ask patients how worried they are wh they contact OOH services, I don't think you can prove with this research that it will (potentially) increase patient safety, you need more evidence for that statement. I have made some specific comments below.</li> <li>Introduction <ol> <li>I think the rationale for this study should be elaborated further in the introduction. Is there evidence that there is a relationship between degree of worry and urgency from a medical perspective Is worry a good indicator for assessing urgency, and thus patient</li> </ol> </li> </ul>

Methods 1. Wasn't it possible to examine the relationship between DOW/ CSM and the urgency assessing according physicians or nurses?
2. Was a triage system used? Did patients themselves mention duration, identity, cause, cure/control and consequence? Or was it asked by the call handler?
3. Why did you exclude children? Especially parents of young children account for a large proportion of the total contacts with OOH services. Worry is an important motive for parents to call. Please explain why you have excluded this group.
4. You mention that the OOH services and the Emergency Medical Services in Copenhagen are integrated. Do you have information about which of the two numbers the included patients have called?
5. Can you explain a bit more about the use of the intensity descriptor. In how many cases did the researchers give a numeric value for DOW?
6. What do you mean by "prior to data collection, two experienced call handlers were asked to assess and recommend revision in question sequence and phrasing in actual calls"? please explain a bit more.
7. Can you mention something about the representativeness of the calls? For example, profession of call handler, how many different call handlers were included, and education level call handlers? Age, gender of patients.
Results 8. In your results, you mention a relationship between duration of the patients' problem and DOW? When I checked figure 3, I can't find this relationship. The medians are similar for the three groups. In general, I don't think you can conclude relationships according this figures/numbers. You need more power. How did you decide when
you can speak of an interrelation? Discussion 9. I think the comparison with other research is too limited. The discussion would be stronger if you discuss your findings more in combination with other literature.
10. Although you mention you use qualitative methods, your analyses about relations are quantitative. Therefore, I think the amount of patients is definitely not a strength. In my opinion, you need more power to make statements about interrelations.
11. What do you mean by "research on optimal patient outcome after DOW-assisted triage is needed"? Please clarify.
Article summary: strength and limitations 12. Your first point is interesting, but you do not describe this point

anywhere else 13. 4th point: "the researcher" should be "one researcher".
13. 4th point. the researcher should be one researcher.

REVIEWER	Joanne Turnbull	
	University of Southampton, UK	
REVIEW RETURNED	19-Jan-2018	
GENERAL COMMENTS	This is a really interesting study with some potentially really interesting findings. However, I don't feel that the findings really do justice to the data that this study probably draws on. The authors suggest that 'This study gives a detailed insight into patients' illness representation in an OOH service' yet all that is presented is some broad themes (derived from a framework) and essentially a frequency count of the number of participants that this theme applied to (see further comments below).	
	1. Abstract: I think the reference to 'strong identity' and 'medium identity' needs some explanation in the abstract. It's not clear what this means from the abstract alone	
	2. Methods: Some explanations of the inclusion/exclusion criteria would be helpful for the reader.	
	3. Methods: The qualitative data analysis description is brief. Given that one of the strengths of the study is said to be 'Use of the NVivo V.11 software and researcher triangulation ensures that the coding of the data is available for independent analysis and less subject to personal bias'. I think a bit more transparency about the coding process / what was actually done is important so that the reader can understand something about the rigour of the analytical process.	
	Results: The paper is presented as a mixed methods study including a qualitative thematic analysis of calls. However, the qualitative findings present little of the data beyond a few quotations in table 1. There is not enough evidence of a detailed / rigorous analysis of the qualitative data. More data presentation is needed to back up the interpretations / assertions that are being made. The paper asserts that 'This study gives a detailed insight into patients' illness representation in an OOH service, which can enable a better understanding of the challenges described in telephone consultations.' However, the paper lacks detail / depth of analysis to support this. The findings / analysis are superficial. There are some broad themes (derived from the framework) and not much more than a count of the responses for each theme e.g. 'possible cause of symptoms or illness was reported by 132 callers (73%)'. The	
	analysis potentially contains some interesting findings but as it currently stands it is not much more than a simple description / frequency count of responses. A greater depth of the textual data / presentation of quotes as evidence, would greatly improve this paper. I would like to have seen more of the types of examples of citations presented in table 1 be a much more prominent part of this paper to really ensure that there is enough data to support the assertions made.	

## VERSION 1 – AUTHOR RESPONSE

## **Reviewer 1**

Reviewer Name: Barbara Farquharson Institution and Country: Edinburgh Napier University, Scotland

Reviewer comments	Author responses	Section reference
The main finding reported is a relationship between illness representations and self-evaluation of urgency (also referred to as degree of worry). However, the theoretical basis of this construct is not described, is measured by a single question and there is no mention of validity and reliability	As this is a new field of study, literature regarding this topic is limited. However, references have been added to the section: <u>Introduction</u> , strengthening the evidence of a relationship between anxiety/DOW and patient- evaluated urgency; and patient-evaluated urgency and patient safety. Furthermore, the question regarding a validated measure for DOW has also been commented on. The revised section: <u>Introduction</u> addresses the above.	Introduction
- is there a validated measure? If not, what data can you offer to support the validity and reliability of this measure?	"Telephone triage within out-of-hours (OOH) service is recognised as a mean to reduce pressure and overcrowding of emergency departments (ED) and OOH clinics. <sup>1</sup> It aims to assess the urgency of a patient's medical condition in order to determine the correct type of health care needed, thus ensuring patient safety. However, due to the lack of non-verbal cues in telephone consultations, assessing urgency is more challenging than face-to-face consultations. <sup>11</sup> Studies show that the quality of telephone triage improves with communication between patient and health professional being patient-centred rather than disease-centred <sup>111</sup> and that non-normative symptom description and poor communication contribute to under- triage. <sup>11</sup> Triage tools, e.g. computerised decision support systems are used to aid the triage process <sup>12</sup> ; however, these tools focus on medical information. <sup>11</sup> Patients' perception of urgency has previously been examined, whereby, ED physicians' and the patients' assessment of the severity of symptoms were compared. <sup>111</sup> These studies found that patients' perception of urgency can be used as a rough guide to predict the need	

for hospitalisation.<sup>ix</sup> Furthermore, it has been suggested that patients expressing a potential need for hospitalisation should be thoroughly examined for possible severe illness.<sup>x</sup> Previous studies have also shown that patient anxiety or worry about a health threat is a major factor in urgent care decision-making.xi xii Therefore, the measure of a patient's anxiety or worry about an acute health threat reflects the patient's selfevaluation of urgency. A self-reported verbal ten-point NRS measuring anxiety in patients (1=minimal anxiety to 10=maximal anxiety) has previously been used in several studies in acute care settings.xiii The anxiety observed in these patients was regarded as acute in relation to the immediate health threat and not due to an underlying psychiatric disease.<sup>xiv</sup> This scale has not been validated. However, as anxiety is a completely subjective symptom, it was felt that a subjective scoring system was acceptable. The feeling of anxiety in this setting is synonymous to worry. In this study, we measured the patient's self-evaluation of urgency, defined as degree-of-worry (DOW), by using a verbal 10-point numerical rating scale (NRS) (1=minimal worry to 10=maximal worry). A previous study shows that callers to OOH services are able to rate their DOW, and that the DOW scale is feasible for large-scale studies.<sup>xv</sup> We used the word worry and not anxiety, as we felt anxiety may be associated with physical symptoms, such as elevated heart rate or shortness of breath, while worry is more of a cognitive nature.xvi The Common-Sense Model of Self-Regulation (CSM), by Leventhal<sup>xvii</sup> is a widely recognized theoretical framework, which can be used to describe how a patient cognitively and emotionally addresses a health threat, based on experienced symptoms. The patient's perception is based on prior experience, personal beliefs, discussions with others and cultural understandings.xviii The CSM is a parallel processing model, with one arm representing the cognitive processing aspects and the other arm representing the emotional processing aspects. Together they make up a

patients' illness representation.<sup>xix</sup> The cognitive

The reference for the intensity descriptor relates to pain - what evidence is there that it is applicable to DOW?	arm can be categorized into five components: 1) identity: symptoms or name/label of the health threat, 2) timeline: duration of the health threat 3) cause: factors that are responsible for the health threat, 4) cure or control! whether the health threat can be cured or controlled and 5) consequence: of the health threat. <sup>xviii</sup> The patients' understanding of their illness representation influences how they present their health issue to a health care provider and this may in turn influence the care they receive. <sup>xx</sup> In previous studies, it has been shown that the five components of the CSM framework account for a large proportion of the presentations patients make when contacting OOH services <sup>voi</sup> and serve as an appropriate framework for understanding the worry experiences of primary health care patients. <sup>xvii</sup> The aim of this paper is to examine the relation between a patient's illness representation, as presented in telephone consultation to an OOH service call handler, and the self-reported DOW as a measure of self-evaluated urgency. If there is a relation, incorporating a patient's DOW as an additional tool in the telephone triage process could aid determination of urgency and type of health care needed, potentially increasing patient safety." In the few calls (n=10), in which the callers failed to give a numerical DOW, the intensity verbal descriptors were applied (see table 1, Duncan <i>et al</i> ). This table lists verbal descriptors rating the intensity of pain and not worry. However, as both pain and worry are subjective, it was felt, that in these few cases, the intensity descriptors for pain were an adequate tool. The section: <u>Data sources</u> has been revised in order to, in more detail, describe the use of the intensity descriptors in the few cases for which intensity descriptors in the few cases for which	Data sources
	these were relevant.	
	"Data consisted of two parallel strands – the quantitative scaled DOW and illness	

representation presented by the callers – both derived from the recorded telephone consultations. Two experienced call handlers were first asked to assess and recommend question phrasing for data collection. All call handlers participated in data collection and received instructions on procedure, inclusion criteria, study focus and voluntary caller participation. Based on the recommendations, call handlers were instructed to ask the following questions in each call: "What is your reason for calling in today?", "How long have you been experiencing these symptoms?" and "On a scale from 1-10, how worried are you?". Additional questions were asked at the call handlers' discretion as they deemed relevant and the caller was invited to participate in the study, giving verbal informed consent. Data were collected throughout the course of the consultation. Calls in which the caller failed to provide a number reflecting their DOW (n=10), were assessed by two researchers and using the intensity verbal descriptors (see table 1, Duncan et al<sup>xxiii</sup>) assigned a numeric value (1 to 10). If not concurrent, a consensus was reached through discussion. The intensity verbal descriptors used, describe the intensity of pain and not worry. However, as both pain and worry are subjective, it was felt, that in these few cases, the intensity descriptors for pain were an adequate tool."

Statements are made about the relationship between constructs (e.g. strong identity and low DOW) and indeed these are the main findings, however they are not supported by any statistics which allow the reader to evaluate the strength and significance of the associations (looking at the graphs in Fig 3, I suspect some would not be significant) - please present the statistical data to support your main findings.	As a new rating scale was used in this study, it was necessary to first train call handlers in using the rating scale in a telephone consultation. It was therefore considered, that a reasonable length of time for using this new rating scale, was three days. This, however, resulted in a limited size of the study population. Due to this limitation, there is a lack of statistical power in this study. However, irrespective of this limitation, the analyses of the results, using the mixed methods approach, show a distinct trend. The results show the relation between DOW as a measure of patient- evaluated urgency and their illness representation. The section: <u>Strengths and limitations of this study</u> is revised to address the above comments.	Strengths and limitations of this study
	"The main strength of this study was the use of mixed methods approach, which gave an in- depth insight and enabled a thorough analysis and understanding of the illness representation of patients to an OOH service. In addition, patients' illness representation and reported self-evaluation of DOW were obtained in real time, as the callers were seeking help. Findings were, therefore, not influenced by recall bias. DOW was not uniformly obtained at a specific time within the consultation. Therefore, the consultation itself could influence the patient's DOW and the patient's DOW could influence the consultation. This, however, is representative of real-life calls to OOH services and how DOW can be used as a potential triage tool. Use of the NVivo V.11 software and researcher triangulation ensure that the coding of the data is available for independent analysis and less subject to personal bias. Due to the short duration of data collection, the size of the study population was limited, resulting in a lack of statistical power. However, irrespective of this limitation, the analyses of the results, using the mixed methods approach, show a distinct trend and relation between DOW as a measure of patient-evaluated urgency and their illness representation."	

It is stated that ethical approval was not required which surprised me - in the UK, a study of this nature would require ethical approval. In any case, details of the steps taken to protect participants/ensure rigor should be described in the manuscript.	The study was approved by the Data Protection Agency and verbal, informed consent from all callers was obtained during the telephone consultation. The sections: <u>Data collection</u> and <u>Data sources</u> (see above) are revised in response to this comment.	Data sources
	" <i>Ethical statement:</i> The study was approved by the Data Protection Agency (PVH-2015-004, I-Suite nr.: 04330). All participants gave informed consent. The Ethical Committee was consulted but no permission was needed (H-15016323)."	Ethical statement
	"A total of 261 callers to the OOH services, The Medical Helpline 1813, during a three-day time period were approached for inclusion in this study. As a new rating scale was being implemented by the call handlers, it was considered, that this was a reasonable length of time. All calls from adults ( $\geq$ 15 years of age) concerning somatic illness were deemed eligible for inclusion. Calls made on behalf of another person, including children (n=16) were excluded, in order to have a study population exclusively describing personal symptoms. Furthermore, calls in which consent was not granted (n=1), calls in which the call handler failed to ask study questions (n=19), calls in which there were technical problems with the call recording (n=33) and repeat callers (n=12) were also excluded. This resulted in a convenience sample of a total of 180 calls. Data were collected for three consecutive days: Wednesday 20 <sup>th</sup> April and Thursday 21 <sup>st</sup> April (4pm to 10pm) and Friday 22 <sup>nd</sup> April (8am to 4pm) 2016 (a bank holiday), 2016. The study was approved by the Data Protection Agency, Denmark."	Data collection
Additional data about the participants would also be helpful to the reader - what age, gender, ethnicity etc,	Callers' age and gender are given in revised table 1 Participant demographics. There is no data on the ethnicity of the callers.	Table 1
-what was the nature of the calls, -were they representative?	The nature of the calls was as follows: acute illness (n = 120), injury (n = 37), exacerbation of chronic disease (n = 15), other (n = 7), and undetermined (n = 1).	Participants

	The nature of the calls was representative of calls to the OOH services. In response to this comment, the section: <u>Participants</u> is revised and a reference (Dam 2017) added. <i>"A total of 261 callers to the OOH services</i>	
	during the three-day time period were approached for inclusion. Of these, 81 callers were excluded, based on the exclusion criteria, leaving a total of 180 callers to be included in this study. Due to this limited size of the study population, there is a lack of statistical power. The nature of the calls was as follows: acute illness ( $n = 120$ ), injury ( $n = 37$ ), exacerbation of chronic disease ( $n = 15$ ), other ( $n = 7$ ), and undetermined ( $n = 1$ ), which is representative for calls to the OOH services. <sup>xxiv</sup> See table 1 for a description of participant demographics."	
-How were the dates for participation chosen?		
	The selected dates for participation fitted with the progression of planning, training of call handlers and the execution of the study.	
For the qualitative analysis - more detail of how coding units were identified would help replicability. More detail as to how blinding to DOW was ensured would provide more reassurance re rigour.	In response to this comment, the section: <u>Qualitative data analysis</u> has been revised to give a more in-depth description of the coding process. During the coding process, blinding for DOW was not carried out, however, DOW was disregarded.	Qualitative data analysis
	"The qualitative data were created by coding the transcripts deductively according to the four components of the CSM framework (identity, cause, cure/control and consequence), while disregarding the DOW value. For each of the four components, data were clustered and patterns identified. Three themes within each component were derived from these patterns and each theme was recoded, as described by Braun and Clarke. <sup>xxv</sup> The patterns and thereby derived theme definitions, were discussed and agreed upon with a second researcher, using	
	50% of the study data. The remaining data were rechecked and recoded if necessary, by	

	the main researcher, according to the agreed theme definitions."	
A limitation identified relates to a single coder having identified the components of CSM - this would be easily overcome by having an additional coder perform a reliability check - providing additional detail of this would increase confidence in the results presented.	The main researcher coded the data and defined themes from the identified patterns, according to the CSM framework. This was then discussed and agreed upon with a second researcher. The section: <u>Qualitative data</u> <u>analysis</u> (see above) has been revised to describe, in more detail, the procedure used. The <u>Contributorship statement</u> is also revised.	Qualitative data analysis
	"Contributorship statement: SLT and HGJ planned the study and discussed and agreed upon theme definitions in the qualitative analysis. HGJ planned and performed the data collection. SLT extracted and analysed the data and drafted the manuscript. IE and FKL supervised and contributed substantially to the critical revision. All authors read and approved the final manuscript."	Contributor ship statement
The results reported relating to timeline describe the duration of symptoms prior to the call - this is of interest but the definition of this construct proposed by Leventhal relates to the perceived duration of the symptoms/condition - can you present additional data which reflects callers perceptions of timeline as well as the duration so far?	Call handlers were instructed to ask each caller: "How long have you been experiencing these symptoms?" as stated in the section: <u>Data sources</u> (see above). The timeline reports are therefore solely based on callers' reports and therefore perception. The information was then grouped into three categories. See <u>table 3 Thematic analysis of</u> the components of the CSM framework In order to improve the richness and quality of the data presented, data was added to the timeline section of <u>table 3 Thematic analysis of</u> the components of the CSM framework. Citations chosen represent the definition of each theme.	Data sources Table 3

The quantitative findings it is reported that callers with a low DOW were more likely to mention a cause for their illness etc. However, as it is not differentiated within the data whether the illness representation was presented spontaneously or prompted by a call-handler, is it not also possible that the relationship could be due to call-handlers being more/less likely to ask prompting questions depending on DOW? Further consideration of this would be helpful.	All call handlers were instructed to ask the following questions in each call: "What is your reason for calling in today?", "How long have you been experiencing these symptoms?" and "On a scale from 1-10, how worried are you?". Additional questions were asked at the call handlers' discretion as they deemed relevant. This resulted in responses being both spontaneous and/or prompted by the call- handler. This is representative of real-life calls to OOH services. In this study DOW was not uniformly obtained at a specific time within the consultation. The consultation itself could influence the patient's DOW and the patient's DOW could influence the consultation. However, this is representative of real-life calls to OOH serves and how DOW can be used as a potential triage tool. This comment is addressed in the revised section: <u>Strengths and limitations of this</u> <u>study</u> (see above).	Data sources Strengths and limitations of this study
A strobe checklist was not available to me - was it submitted?	A STROBE checklist is completed and included in the resubmission.	Strobe checklist
Minor point but might it be preferable to phrase things in a non-gendered way (e.g. 'their illness representations' rather than 'her illness representations')	The word "her" is replaced by "their".	Throughout the paper

Würgler MW, Navne LE. Når sygeplejersker visiterer i lægevagten. Copenhagen: Dansk Sundhedsinstitut; 2010.

Leprohon J, Patel VL. Decision-making Strategies for Telephone Triage in Emergency Medical Services. Medical Decision Making. 1995;15:240–53. doi:10.1177/0272989x9501500307

Derkx HP. For your ears only. Quality of telephone triage at out-of-hours centres in the Netherlands. Maastricht: Department of General Practice, University of Maastricht; 2008.

Gamst-Jensen H, Lippert FK, Egerod I. Under-triage in telephone consultation is related to nonnormative symptom description and interpersonal communication: a mixed methods study. Scandinavian Journal of Trauma, Resuscitation and Emergency Medi-cine. 2017;25:52. doi:10.1186/s13049-017-0390-0 Gravsersen DS. Out-of.hours telephone triage by nurses and doctors in Danish acute care settings – A study of quality focusing on communication, safety and efficiency. Aarhus: Re-search Unit for General Practice, University of Aarhus; 2016.

Holmström I. Decision aid software programs in telenursing: not used as intended? Experi-ences of Swedish telenurses. Nursing & Health Sciences. 2007;9:23–8. doi:10.1111/j.1442-2018.2007.00299.x

Gifford MJ, Franaszek JB, Gibson G. Emergency physicians and patients assessments: Urgency of need for medical care. Annals of Emergency Medicine. 1980;9:502–7. doi:10.1016/s0196-0644(80)80187-9

Hunt RC, Dehart KL, Allison E, et al. Patient and physician perception of need for emer-gency medical care: A prospective and retrospective analysis. The American Journal of Emergency Medicine. 1996;14:635–9. doi:10.1016/s0735-6757(96)90077-7

Caterino JM, Holliman C, Kunselman AR. Underestimation of case severity by emergency department patients: Implications for managed care. The American Journal of Emergency Medicine. 2000;18:254–6. doi:10.1016/s0735-6757(00)90115-3

Miyamichi R, Mayumi T, Asaoka M, et al. Evaluating patient self-assessment of health as a predictor of hospital admission in emergency practice: a diagnostic validity study. Emergency Medicine Journal. 2011;29:570–5. doi:10.1136/emj.2010.105247

Booker MJ, Simmonds RL, Purdy S. Patients who call emergency ambulances for primary care problems: a qualitative study of the decision-making process. Emergency Medicine Journal. 2013;31:448–52. doi:10.1136/emermed-2012-202124

Agarwal S, Banerjee J, Baker R, et al. Potentially avoidable emergency department at-tendance: interview study of patients reasons for attendance. Emergency Medicine Jour-nal 2011;29. doi:10.1136/emermed-2011-200585

Fosnocht D, Swanson E. Pain and anxiety in the emergency department. Annals of Emer-gency Medicine. 2004;44. doi:10.1016/j.annemergmed.2004.07.284

Craven P, Cinar O, Madsen T. Patient anxiety may influence the efficacy of ED pain man-agement. The American Journal of Emergency Medicine. 2013;31:313–8. doi:10.1016/j.ajem.2012.08.009

Gamst-Jensen H, Huibers L, Pedersen K, et al. Self-rated worry in acute care telephone triage: a mixedmethods study. British Journal of General Practice Published Online First: December 2018. doi:10.3399/bjgp18x695021

Zebb BJ, Beck JG. Worry Versus Anxiety. Behavior Modification. 1998;22:45–61. doi:10.1177/01454455980221003

Leventhal H, Meyer D, Nerenz D. The common sense representation of illness danger. In: Rachman S (ed). Contributions to medical psychology. 1st edn,Vol 2.Oxford: Pergamon Press, 1980:7–30.

Leventhal H, Phillips LA, Burns E. The Common-Sense Model of Self-Regulation (CSM): a dynamic framework for understanding illness self-management. Journal of Behavioral Medicine. 2016;39:935–46. doi:10.1007/s10865-016-9782-2

Hagger MS, Orbell S. A Meta-Analytic Review of the Common-Sense Model of Illness Representations. Psychology & Health. 2003;18:141–84. doi:10.1080/088704403100081321

Bugge C, Entwistle VA, Watt IS. The significance for decision-making of information that is not exchanged by patients and health professionals during consultations. Social Science & Medicine. 2006;63:2065–78. doi:10.1016/j.socscimed.2006.05.010

Farquharson B, Johnston M, Bugge C. How people present symptoms to health services: a theorybased content analysis. British Journal of General Practice. 2011;61:267–73. doi:10.3399/bjgp11x567090 Laakso V, Niemi PM. Primary health-care patients' reasons for complaint-related worry and relief.

Primary Health Care Research & Development. 2013;14:151–63. doi:10.1017/s1463423612000448

Duncan GH, Bushnell CM, Lavigne GJ. Comparison of verbal and visual analogue scales for measuring the intensity and unpleasantness of experimental pain. Pain. 1989;37:295–303. doi:10.1016/0304-3959(89)90194-2

Dam, PS. Derfor ringer danskerne 1813. Berlingske □newspaper on the internet □. 2017 Mar 19 □cited 2018 Feb 10 □. Available from: https://www.b.dk/nationalt/derfor-ringer-danskerne-1813

Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psy-chology. 2006;3:77–101. doi:10.1191/1478088706qp063oa

#### **Reviewer 2**

Reviewer Name: Ellen Keizer

Institution and Country: Institute of Primary Care, University Hospital Zürich, Zürich, Switzerland

Reviewer comments	Author response	Section reference
Introduction		
I think the rationale for this study should be elaborated further in the introduction. Is there evidence that there is a relationship between degree of worry and urgency from a medical perspective?	As this is a new field of study, literature regarding this topic is limited. However, references have been added to the section: <u>Introduction</u> , strengthening the evidence of a relationship between anxiety/DOW and patient- evaluated urgency; and patient-evaluated urgency and patient safety. The revised section: <u>Introduction</u> addresses the above.	Introduction
Is worry a good indicator for assessing urgency, and thus patient safety? I think this statement need some evidence. You should refer to studies which have found a relationship between worry and patient safety (or	"Telephone triage within out-of-hours (OOH) service is recognised as a mean to reduce pressure and overcrowding of emergency departments (ED) and OOH clinics. <sup>xxvi</sup> It aims to assess the urgency of a patient's medical condition in order to determine the correct type	

	1	
between self-evaluated urgency and patient safety).	of health care needed, thus ensuring patient safety. However, due to the lack of non-verbal cues in telephone consultations, assessing urgency is more challenging than face-to-face consultations. <sup>xxvii</sup> Studies show that the quality of telephone triage improves with communication between patient and health professional being patient-centred rather than disease-centred <sup>xxvii</sup> and that non-normative symptom description and poor communication contribute to under-triage. <sup>xxix</sup> Triage tools, e.g. computerised decision support systems are used to aid the triage process <sup>xxx</sup> ; however, these tools focus on medical information and less on psychosocial or affective information. <sup>xxxi</sup>	
	less on psychosocial or affective information. <sup>xxxi</sup> Patients' perception of urgency has previously been examined, whereby, ED physicians' and the patients' assessment of the severity of symptoms were compared. <sup>xxxii xxxiii</sup> These studies found that patients' perception of urgency can be used as a rough guide to predict the need for hospitalisation. <sup>xxxiv</sup> Furthermore, it has been suggested that patients expressing a potential need for hospitalisation should be thoroughly examined for possible severe illness. <sup>xxxv</sup> Previous studies have also shown that patient anxiety or worry about a health threat is a major factor in urgent care decision-making. <sup>xxxvi xxxvii</sup> Therefore, the measure of a patient's anxiety or worry about an acute health threat reflects the patient's self- evaluation of urgency. A self-reported verbal ten-point NRS measuring anxiety in patients (1=minimal anxiety to 10=maximal anxiety) has previously been used in several studies in acute care settings. <sup>xxxviiii</sup> The anxiety observed in these patients was regarded as acute in relation to the immediate health threat and not due to an underlying psychiatric disease. <sup>xxxix</sup> This scale has not been validated. However, as	
	anxiety is a completely subjective symptom, it was felt that a subjective scoring system was acceptable. The feeling of anxiety in this setting is synonymous to worry. In this study, we measured the patient's self-evaluation of urgency, defined as degree-of-worry (DOW), by using a verbal 10-point numerical rating scale (NRS) (1=minimal worry to 10=maximal worry).	

A previous study shows that callers to OOH services are able to rate their DOW, and that the DOW scale is feasible for large-scale studies. <sup>xl</sup> We used the word worry and not anxiety, as we felt anxiety may be associated with physical symptoms, such as elevated heart rate or shortness of breath, while worry is more of a cognitive nature. <sup>xli</sup>	
The Common-Sense Model of Self-Regulation (CSM), by Leventhal <sup>x/iii</sup> is a widely recognized theoretical framework, which can be used to describe how a patient cognitively and emotionally addresses a health threat, based on experienced symptoms. The patient's perception is based on prior experience, personal beliefs, discussions with others and cultural understandings. <sup>x/iiii</sup> The CSM is a parallel processing model, with one arm representing the cognitive processing aspects	
and the other arm representing the emotional processing aspects. Together they make up a patients' illness representation. <sup>xliv</sup> The cognitive arm can be categorized into five components: 1) identity: symptoms or name/label of the health threat, 2) timeline: duration of the health threat 3) cause: factors that are responsible for the health threat, 4) cure or control: whether the health threat can be cured or controlled and 5) consequence: of the health threat. <sup>xviii</sup> The patients' understanding of their illness	
representation influences how they present their health issue to a health care provider and this may in turn influence the care they receive. <sup>xlv</sup> In previous studies, it has been shown that the five components of the CSM framework account for a large proportion of the presentations patients make when contacting OOH services <sup>xlvi</sup> and serve as an appropriate framework for understanding the worry experiences of primary health care patients. <sup>xlvii</sup>	
The aim of this paper is to examine the relation between a patient's illness representation, as presented in telephone consultation to an OOH service call handler, and the self-reported DOW as a measure of self-evaluated urgency. If there is a relation, incorporating a patient's DOW as an additional tool in the telephone triage process could aid determination of	

	urgency and type of health care needed, potentially increasing patient safety."	
Methods		
Wasn't it possible to examine the relationship between DOW/ CSM and the urgency assessing according physicians or nurses?	In the present study, it was not possible to examine the relationship between patient-rated DOW as a measure of perception of urgency and the call handlers' assessment of urgency. However, future research on this relationship both prospectively and retrospectively is presently being discussed. The paper referenced in the introduction: Gifford MJ, Franaszek JB, Gibson G. Emergency physicians and patients assessments: Urgency of need for medical care. Annals of Emergency Medicine 1980;9:502–7. doi:10.1016/s0196- 0644(80)80187-9 can be used as inspiration for further research. This is addressed in the revised section: Relevance of this study: possible implications for health care providers and policy makers. "This study suggests a relation between a patient's illness representation, as presented telephonically to an OOH services call handler, and the self-evaluation of urgency, defined as DOW. This is a new area of research and this study gives direction for future research to further strengthen the evidence. Research on coherence between patient DOW, call handlers', ED and GP physicians' assessment of urgency, both prospectively and retrospectively will strengthen the basis for potential use of DOW as a triage tool. Incorporating DOW as an additional tool in the telephone triage process could potentially aid in the determination of urgency and the type of health care needed, thus increasing patient safety."	Relevance of this study: possible implications for health care providers and policy makers

Was a triage system used?	Call handlers in the OOH services in Copenhagen, Denmark use a non-validated triage system, inspired by a system used in Sweden and a criteria-based dispatch protocol termed the Danish Index for Emergency Care.	
	The call handlers' clinical decision making is guided by a locally developed criterion-based decision tool.	
Did patients themselves mention		Data sources
duration, identity, cause, cure/control and consequence? Or was it asked by the call handler?	Call handlers were instructed to ask the following questions in each call: "What is your reason for calling in today?", "How long have you been experiencing these symptoms?" and "On a scale from 1-10, how worried are you?". Additional questions were asked at the call handlers' discretion as they deemed relevant. This resulted in responses being both spontaneous and/or prompted by the call- handler. This is representative of real-life calls to OOH services. This is addressed in the revised section: <u>Data sources</u> .	
	"Data consisted of two parallel strands – the quantitative scaled DOW and illness representation presented by the callers – both derived from the recorded telephone consultations. Two experienced call handlers were first asked to assess and recommend question phrasing for data collection. All call	
	handlers participated in data collection and received instructions on procedure, inclusion criteria, study focus and voluntary caller participation. Based on the recommendations,	
	call handlers were instructed to ask the following questions in each call: "What is your reason for calling in today?", "How long have	
	you been experiencing these symptoms?" and "On a scale from 1-10, how worried are you?". Additional questions were asked at the call	
	handlers' discretion as they deemed relevant and the caller was invited to participate in the	
	study, giving verbal informed consent. Data were collected throughout the course of the	
	consultation. Calls in which the caller failed to	
	provide a number reflecting their DOW (n=10), were assessed by two researchers and using	
	the intensity verbal descriptors (see table 1,	

Why did you exclude children? Especially parents of young children account for a large proportion of the total contacts with OOH services. Worry is an important motive for parents to call. Please explain why you have excluded this group.	In this study, we wished to focus on the scaled DOW from the actual patient and not a proxy, thereby, having a study population exclusively describing personal symptoms. Therefore, all calls made on behalf of another person including children under the age of 15 years were excluded, as stated in the section: <u>Data</u> <u>collection</u> .	Data collection
	We acknowledge that parents' DOW on behalf of their children is an important aspect of consultations in OOH services. However, as this is the first study examining patients' DOW in this setting, we made the above-mentioned decision. Incorporating parents' DOW on behalf of their children is a potential area for future research.	
	"A total of 261 callers to the OOH services, The Medical Helpline 1813, during a three-day time period were approached for inclusion in this study. As a new rating scale was being implemented by the call handlers, it was considered, that this was a reasonable length of time. All calls from adults ( $\geq$ 15 years of age) concerning somatic illness were deemed eligible for inclusion. Calls made on behalf of another person, including children (n=16) were excluded, in order to have a study population exclusively describing personal symptoms. Furthermore, calls in which consent was not granted (n=1), calls in which the call handler failed to ask study questions (n=19), calls in	

	which there were technical problems with the call recording (n=33) and repeat callers (n=12) were also excluded. This resulted in a convenience sample of a total of 180 calls. Data were collected for three consecutive days: Wednesday 20 <sup>th</sup> April and Thursday 21 <sup>st</sup> April (4pm to 10pm) and Friday 22 <sup>nd</sup> April (8am to 4pm) 2016 (a bank holiday), 2016. The study was approved by the Data Protection Agency, Denmark."	
You mention that the OOH services and the Emergency Medical Services in Copenhagen are integrated. Do you have information about which of the two numbers the included patients have called?	Only calls to the The Medical Helpline 1813 were included in this study. Calls to this number comprise of non-emergent medical calls as opposed to 112 for life-threatening emergencies. We only included calls to The Medical Helpline 1813, as all calls concerning life-threatening problems were excluded, as stated in the section: <u>Data collection</u> (see above). The section: <u>Setting</u> has been revised for greater precision.	Data collection Setting
	"The OOH services and the Emergency Medical Services, Copenhagen, the Capital Region of Denmark, are integrated in one organisation and can be reached through two telephone numbers; 112 for life-threatening emergencies and 1813 for acute, non- emergent medical calls. The Medical Helpline 1813 is available from 4pm to 8am on weekdays and around the clock on weekends and holidays. Individuals may also call 1813 for a referral to an emergency department, if they cannot get in touch with their general practice (GP) during regular working hours. All access to acute care is pre-assessed by telephone triage. Annually, approximately one million calls are handled by call handlers (nurses/physicians) who triage the caller to self-care, a general practitioner, face-to-face assessment/consultation at a hospital, home visit or direct hospitalisation. <sup>xlix Im</sup>	
Can you explain a bit more about the use of the intensity descriptor. In how many cases did the researchers give a numeric value for DOW?	In the few calls (n=10), in which the callers failed to give a numerical DOW, the intensity verbal descriptors were applied (see table 1, Duncan <i>et al</i> ). This table lists verbal descriptors rating the intensity of pain and not worry. However, as both pain and worry are subjective, it was felt, that in these few cases,	Data sources

	the intensity descriptors for pain were an adequate tool. The section: <u>Data sources</u> (see above) has been revised in order to, in more detail, describe the use of the intensity descriptor in the few cases for which this was relevant.	
What do you mean by "prior to data collection, two experienced call handlers were asked to assess and recommend revision in question sequence and phrasing in actual calls"? please explain a bit more.	The recommendations given by the two experienced call handlers prior to data collection were phrasing of call handlers' questions such as: "What is your reason for calling in today?", "How long have you been experiencing these symptoms?" and "On a scale from 1-10, how worried are you?". Furthermore, they recommended asking questions regarding the duration of symptoms before asking the patient to rate worry, as this is in line with the natural history taking flow. The section: <u>Data sources</u> (see above) has been revised for greater precision.	Data sources

Can you mention something about the representativeness of the calls?	The nature of the calls was as follows: acute illness (n=120), injury (n=37), exacerbation of chronic disease (n=15), other (n=7), and undetermined (n=1). These were representative of calls to the OOH services. In response to this comment, the section: <u>Participants</u> is revised and a reference (Dam 2017) added.	Participants
	"A total of 261 callers to the OOH services during the three-day time period were approached for inclusion. Of these, 81 callers were excluded, based on the exclusion criteria, leaving a total of 180 callers to be included in this study. Due to this limited size of the study population, there is a lack of statistical power. The nature of the calls was as follows: acute illness ( $n = 120$ ), injury ( $n = 37$ ), exacerbation of chronic disease ( $n = 15$ ), other ( $n = 7$ ), and undetermined ( $n = 1$ ), which is representative for calls to the OOH services. <sup>II</sup> See table 1 for a description of participant demographics."	
For example, profession of call handler, how many different call handlers were included, and education level call handlers?	Unfortunately, we are not in possession of information regarding the profession or education of the call handlers. However, in order to be employed as call handler at the OOH services, nurses need at least five years of experience and physicians need to be specialized or in training to become a specialist. We are not in possession of information regarding the number of caller handlers. However, up to 45 call handlers can be present during a shift.	
Age, gender of patients.	Callers' age and gender are given in revised table 1 Participant demographics.	

	Table 1
Results	

In your results, you mention a relationship between duration of the patients' problem and DOW? When I checked figure 3, I can't find this relationship. The medians are similar for the three groups.	In figure 2 Relation between DOW and the components of the CSM framework, the box for timeline shows similar medians. However, it also shows that the DOW rating for callers who described symptoms which had lasted more than 24 hours where in a higher range than those from callers who described symptoms which had lasted less than 5 hours. As described in the section: Mixed methods findings. <i>"Callers whose illness had lasted less than 5 hours were more likely to have a low to moderate DOW, whereas callers whose illness had lasted more likely to have a more moderate to high DOW."</i>	Mixed methods findings
In general, I don't think you can conclude relationships according this figures/numbers. You need more power. How did you decide when you can speak of an interrelation?	As a new rating scale was used in this study, it was necessary to first train call handlers in using the rating scale in a telephone consultation. It was therefore considered, that a reasonable length of time for using this new rating scale, was three days. This, however, resulted in a limited size of the study population. Due to this limitation, there is a lack of statistical power in this study. However, irrespective of this limitation, the analyses of the results, using the mixed methods approach, show a distinct trend. The results show the relation between DOW as a measure of patient-evaluated urgency and their illness representation.	
	The word "interrelation" is replaced with " <i>relation</i> ". This change addresses the comment made on the lack of statistical power in this study.	Title
		Abstract
		Throughout the paper

Discussion		
I think the comparison with other research is too limited. The discussion would be stronger if you discuss your findings more in combination with other literature.	With respect to the discussion of this study and comparison to other research; as this is a relative new field of research, existing literature is limited. However, the theoretical basis for this study has been elaborated on in the section: <u>Introduction</u> (see above), incorporating additional relevant literature.	Introduction
Although you mention you use qualitative methods, your analyses about relations are quantitative. Therefore, I think the amount of patients is definitely not a strength. In my opinion, you need more power to make statements about interrelations.	The word "interrelation" is replaced with " <i>relation</i> ". This change addresses the comment made on the lack of statistical power in this study. However, irrespective of this limitation, the analyses of the results, using the mixed methods approach, show a distinct trend. The results show the relation between DOW as a measure of patient-evaluated urgency and their illness representation.	Title Abstract Throughout the paper
What do you mean by "research on optimal patient outcome after DOW- assisted triage is needed"? Please clarify.	In the present study, it was not possible to examine the relationship between patient-rated DOW as a measure of perception of urgency and the call handlers' assessment of urgency. However, future research on this relationship both prospectively and retrospectively is presently being discussed. The paper referenced in the introduction: Gifford MJ, Franaszek JB, Gibson G. Emergency physicians and patients assessments: Urgency of need for medical care. <i>Annals of Emergency</i> <i>Medicine</i> 1980; <b>9</b> :502–7. doi:10.1016/s0196- 0644(80)80187-9 can be used as inspiration for further research. This is addressed in the section: <u>Relevance of this study: possible implications for health care providers and policy makers</u> (see above).	Relevance of this study: possible implications for health care providers and policy makers
Article summary: strength and limitations		
Your first point is interesting, but you do not describe this point anywhere else	In response to this comment, the first point is revised to more specifically describe this study and states: "Use of mixed methods approach in this study gave an in-depth insight and enabled a thorough analysis and understanding of the	Article summary: Strengths and limitations

	illness representation of patients to an OOH service."	of this study
4th point: "the researcher" should be "one researcher".	This revision has been made: "the" is replaced with " <i>one</i> "	Article summary: Strengths and limitations of this study

Würgler MW, Navne LE. Når sygeplejersker visiterer i lægevagten. Copenhagen: Dansk Sundhedsinstitut; 2010.

Leprohon J, Patel VL. Decision-making Strategies for Telephone Triage in Emergency Medical Services. Medical Decision Making. 1995;15:240–53. doi:10.1177/0272989x9501500307

Derkx HP. For your ears only. Quality of telephone triage at out-of-hours centres in the Netherlands. Maastricht: Department of General Practice, University of Maastricht; 2008.

Gamst-Jensen H, Lippert FK, Egerod I. Under-triage in telephone consultation is related to nonnormative symptom description and interpersonal communication: a mixed methods study. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine. 2017;25:52. doi:10.1186/s13049-017-0390-0

Gravsersen DS. Out-of.hours telephone triage by nurses and doctors in Danish acute care settings – A study of quality focusing on communication, safety and efficiency. Aarhus: Research Unit for General Practice, University of Aarhus; 2016.

Holmström I. Decision aid software programs in telenursing: not used as intended? Experiences of Swedish telenurses. Nursing & Health Sciences. 2007;9:23–8. doi:10.1111/j.1442-2018.2007.00299.x

Gifford MJ, Franaszek JB, Gibson G. Emergency physicians and patients assessments: Urgency of need for medical care. Annals of Emergency Medicine. 1980;9:502–7. doi:10.1016/s0196-0644(80)80187-9

Hunt RC, Dehart KL, Allison E, et al. Patient and physician perception of need for emergency medical care: A prospective and retrospective analysis. The American Journal of Emergency Medicine. 1996;14:635–9. doi:10.1016/s0735-6757(96)90077-7

Caterino JM, Holliman C, Kunselman AR. Underestimation of case severity by emergency department patients: Implications for managed care. The American Journal of Emergency Medicine. 2000;18:254–6. doi:10.1016/s0735-6757(00)90115-3

Miyamichi R, Mayumi T, Asaoka M, et al. Evaluating patient self-assessment of health as a predictor of hospital admission in emergency practice: a diagnostic validity study. Emergency Medicine Journal. 2011;29:570–5. doi:10.1136/emj.2010.105247

Booker MJ, Simmonds RL, Purdy S. Patients who call emergency ambulances for primary care problems: a qualitative study of the decision-making process. Emergency Medicine Journal. 2013;31:448–52. doi:10.1136/emermed-2012-202124

Agarwal S, Banerjee J, Baker R, et al. Potentially avoidable emergency department attendance: interview study of patients reasons for attendance. Emergency Medicine Journal 2011;29. doi:10.1136/emermed-2011-200585

Fosnocht D, Swanson E. Pain and anxiety in the emergency department. Annals of Emergency Medicine. 2004;44. doi:10.1016/j.annemergmed.2004.07.284

Craven P, Cinar O, Madsen T. Patient anxiety may influence the efficacy of ED pain management. The American Journal of Emergency Medicine. 2013;31:313–8. doi:10.1016/j.ajem.2012.08.009

Gamst-Jensen H, Huibers L, Pedersen K, et al. Self-rated worry in acute care telephone triage: a mixedmethods study. British Journal of General Practice Published Online First: December 2018. doi:10.3399/bjgp18x695021

Zebb BJ, Beck JG. Worry Versus Anxiety. Behavior Modification. 1998;22:45–61. doi:10.1177/01454455980221003

Leventhal H, Meyer D, Nerenz D. The common sense representation of illness danger. In: Rachman S (ed). Contributions to medical psychology. 1st edn,Vol 2.Oxford: Pergamon Press, 1980:7–30.

Leventhal H, Phillips LA, Burns E. The Common-Sense Model of Self-Regulation (CSM): a dynamic framework for understanding illness self-management. Journal of Behavioral Medicine. 2016;39:935–46. doi:10.1007/s10865-016-9782-2

Hagger MS, Orbell S. A Meta-Analytic Review of the Common-Sense Model of Illness Representations. Psychology & Health. 2003;18:141–84. doi:10.1080/088704403100081321

Bugge C, Entwistle VA, Watt IS. The significance for decision-making of information that is not exchanged by patients and health professionals during consultations. Social Science & Medicine. 2006;63:2065–78. doi:10.1016/j.socscimed.2006.05.010

Farquharson B, Johnston M, Bugge C. How people present symptoms to health services: a theorybased content analysis. British Journal of General Practice. 2011;61:267–73. doi:10.3399/bjgp11x567090 Laakso V, Niemi PM. Primary health-care patients' reasons for complaint-related worry and relief.

Primary Health Care Research & Development. 2013;14:151–63. doi:10.1017/s1463423612000448 Duncan GH, Bushnell CM, Lavigne GJ. Comparison of verbal and visual analogue scales for measuring the intensity and unpleasantness of experimental pain. Pain. 1989;37:295–303. doi:10.1016/0304-3959(89)90194-2

Årsrapportdata. Copenhagen: Emergency Medical Services Copenhagen, University of Copenhagen; 2015 and 2016. [cited 2017 Apr 15]. Available from: https://www.regionh.dk/om-region-hovedstaden/Den-Praehospitale-Virksomhed/Documents/%C3%85rsrapportdata20152016.pdf

Forde I, Nader C, Socha-Dietrich K, Oderkirk J, Colombo F. Primary Care Review of Denmark. OECD; 2016. [cited 2017 Apr 11]. Available from: https://www.oecd.org/health/health-systems/Primary-Care-Review-of-Denmark-OECD-report-December-2016.pdf

Dam, PS. Derfor ringer danskerne 1813. Berlingske 

newspaper on the internet

. 2017 Mar 19

cited 2018 Feb 10
. Available from: <a href="https://www.b.dk/nationalt/derfor-ringer-danskerne-1813">https://www.b.dk/nationalt/derfor-ringer-danskerne-1813</a>

## **Reviewer 3**

Reviewer Name: Joanne Turnbull Institution and Country: University of Southampton, UK

Reviewer comments	Author responses	Section reference
Abstract		
I think the reference to 'strong identity' and 'medium identity' needs some explanation in the abstract. It's not clear what this means from the abstract alone	In response to this comment, the section: <u>Results</u> of the <u>Abstract</u> has been revised to include the sentence: " <i>Through</i> <i>qualitative analysis, themes were</i> <i>defined.</i> " As the abstract is limited in length, and the definitions of the themes are comprehensive, this sentence is used to refer to the definitions of the themes given in the section: <u>Qualitative findings</u> in the manuscript.	Abstract Qualitative findings
Methods		
Some explanations of the inclusion/exclusion criteria would be helpful for the reader.	In response to this comment, data from figure 1: <u>Flowchart of calls included</u> , have been incorporated into the section: <u>Data</u> <u>collection</u> , giving a more in-depth description of the inclusion/exclusion criteria. Figure 1 is deleted in the revised manuscript and replaced by table 1: <u>Participant demographics</u> . This revision is due to inquiries from the two other reviewers concerning participant demographics.	Data collection
	"A total of 261 callers to the OOH services, The Medical Helpline 1813, during a three-day time period were	

	approached for inclusion in this study. As a new rating scale was being implemented by the call handlers, it was considered, that this was a reasonable length of time. All calls from adults ( $\geq$ 15 years of age) concerning somatic illness were deemed eligible for inclusion. Calls made on behalf of another person, including children (n=16) were excluded, in order to have a study population exclusively describing personal symptoms. Furthermore, calls in which consent was not granted (n=1), calls in which the call handler failed to ask study questions (n=19), calls in which there were technical problems with the call recording (n=33) and repeat callers (n=12) were also excluded. This resulted in a convenience sample of a total of 180 calls. Data were collected for three consecutive days: Wednesday 20 <sup>th</sup> April and Thursday 21 <sup>st</sup> April (4pm to 10pm) and Friday 22 <sup>nd</sup> April (8am to 4pm) 2016 (a bank holiday), 2016. The study was approved by the Data Protection Agency, Denmark."	
The qualitative data analysis description is brief. Given that one of the strengths of the study is said to be 'Use of the NVivo V.11 software and researcher triangulation ensures that the coding of the data is available for independent analysis and less subject to personal bias'. I think a bit more transparency about the coding process / what was actually done is important so that the reader can understand something about the rigour of the analytical process.	In response to this comment, the section: <u>Qualitative data analysis</u> is revised to give a more in-depth description of the coding process. The main researcher coded the data and defined themes from the identified patterns, according to the CSM framework. This was then discussed and agreed upon with a second researcher. The <u>Contributorship statement</u> is also revised.	Qualitative data analysis
	"The qualitative data were created by coding the transcripts deductively according to the four components of the CSM framework (identity, cause, cure/control and consequence), while disregarding the DOW value. For each of the four components, data were clustered and patterns identified. Three themes	

	within each component were derived from these patterns and each theme was recoded, as described by Braun and Clarke. <sup>III</sup> The patterns and thereby derived theme definitions, were discussed and agreed upon with a second researcher, using 50% of the study data. The remaining data were rechecked and recoded if necessary, by the main researcher, according to the agreed theme definitions." <b>"Contributorship statement:</b> SLT and HGJ planned the study and discussed and agreed upon theme definitions in the qualitative analysis. HGJ planned and performed the data collection. SLT extracted and analysed the data and drafted the manuscript. IE and FKL supervised and contributed substantially to the critical revision. All authors read and approved the final manuscript."	Contributorship statement
Results		
The paper is presented as a mixed methods study including a qualitative thematic analysis of calls. However, the qualitative findings present little of the data beyond a few quotations in table 1. There is not enough evidence of a detailed / rigorous analysis of the qualitative data. More data presentation is needed to back up the interpretations / assertions that are being made.	In response to this comment, the section: <u>Qualitative data analysis (</u> see above) is revised to give a more in-depth description of the coding process. Furthermore, in order to improve the richness and quality of the data presented, more data was added to table 3. titled: <u>Thematic analysis of the</u> <u>components of the CSM framework</u> . Citations chosen represent the breadth of the definition of each theme.	Qualitative data analysis Table 3
The paper asserts that 'This study gives a detailed insight into patients' illness representation in an OOH service, which can enable a better understanding of the challenges described in telephone consultations.' However, the paper lacks detail / depth of analysis to support this.	In response to this comment, the statement: "This study gives a detailed insight into patients' illness representation in OOH service, which can enable a better understanding of the challenges described	

	more specifically describe this study and states: "Use of mixed methods approach in this study gave an in-depth insight and enabled a thorough analysis and understanding of the illness representation of patients to an OOH service."	Article summary: Strengths and limitations of this study
The findings / analysis are superficial. There are some broad themes (derived from the framework) and not much more than a count of the responses for each theme e.g. 'possible cause of symptoms or illness was reported by 132 callers (73%)'. The analysis potentially contains some interesting findings but as it currently stands it is not much more than a simple description / frequency count of responses.	As a new rating scale was used in this study, it was necessary to first train call handlers in using the rating scale in a telephone consultation. It was therefore considered, that a reasonable length of time for using this new rating scale, was three days. This, however, resulted in a limited size of the study population. Due to this limitation, there is a lack of statistical power in this study. However, irrespective of this limitation, the analyses of the results, using the mixed methods approach, show a distinct trend. The results show the relation between DOW as a measure of patient-evaluated urgency and their illness representation. This is addressed in the revised section: <u>Strengths and limitations of this study.</u>	Strengths and limitations of this study
	"The main strength of this study was the use of mixed methods approach, which gave an in-depth insight and enabled a thorough analysis and understanding of the illness representation of patients to an OOH service. In addition, patients' illness representation and reported self- evaluation of DOW were obtained in real time, as the callers were seeking help. Findings were, therefore, not influenced by recall bias. DOW was not uniformly obtained at a specific time within the consultation. Therefore, the consultation itself could influence the patient's DOW and the patient's DOW could influence the consultation. This, however, is representative of real-life calls to OOH services and how DOW can be used as a	

	potential triage tool. Use of the NVivo V.11 software and researcher triangulation ensure that the coding of the data is available for independent analysis and less subject to personal bias. Due to the short duration of data collection, the size of the study population was limited, resulting in a lack of statistical power. However, irrespective of this limitation, the analyses of the results, using the mixed methods approach, show a distinct trend and relation between DOW as a measure of patient-evaluated urgency and their illness representation."	
A greater depth of the textual data / presentation of quotes as evidence, would greatly improve this paper. I would like to have seen more of the types of examples of citations presented in table 1 be a much more prominent part of this paper to really ensure that there is enough data to support the assertions made.	In response to this comment, more data are added to <u>table 3 Thematic analysis of</u> <u>the components of the CSM framework</u> , in order to improve the richness and quality of the data presented. Citations chosen represent the breadth of the definition of each theme.	Table 3

Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology. 2006;3:77–101. doi:10.1191/1478088706qp063oa

## **VERSION 2 – REVIEW**

Ellen Keizer
Institute for Primary Care, University Zürich, Switzerland
21-Mar-2018

GENERAL COMMENTS	General
	I think the authors have done a good job in improving the manuscript. However, I still have doubts about the usefulness and the quality of this study.
	In my view, the conclusion is that patients who are insecure about the reason of their problem (week identity, unclear cause, unclear solution) are more worried than patients who do have an idea. I

think you should more focus on this in your conclusions. Why is this interesting to know?
Before you can say something about safety, you definitely should measure the urgency estimation of medical professionals (you have mentioned this).
Besides, I think the methods of this study are weak. Finding relations indicates testing, but this has not happened. Replacing the word 'interrelation' with 'relation' does not help. In addition, the qualitative analysis seems to be poor. The only thing you do is counting things. So, I do not agree with the terms 'in-depth insight and 'thorough analysis', which you mention in the strengths and limitations. But I am not an expert in qualitative analyses; perhaps an expert could be consulted.
Introduction:
In think the information about what you did in this study (in this study, we measured) fits better in the methods.
Main findings
I cannot conclude from figure 2 that there is a relation between an illness duration of less than five hours and a low DOW. Besides, I do not think it is a main finding that "all callers referred to timeline". The call handler asked them.
Since you did not test the relation between the five components and self-evaluated DOW, I think you should be more careful in the formulation of your conclusions. I would prefer something to say like: it seems that (a low DOW was more present)
Comparison with existing literature
I still think you should compare your results more with other literature.

# VERSION 2 – AUTHOR RESPONSE

Institution and Country: Institute of Primary Care, University Hospital Zürich, Zürich, Switzerland

Reviewer comments	Author response	Section reference
General		
I think the authors have done a good job in improving the manuscript. However, I still have doubts about the usefulness and the quality of this study.	We believe this study is essential, as it shows a relation between patients' illness representation and their DOW. The relation observed, is that DOW is not random, but follows a pattern, depending on the caller's illness representation. This provides an insight in patients' DOW in relation to their illness representation, which can aid call handlers in understanding callers' perception of urgency, potentially improving the triage process.	Relevance of this study: possible implications for health care providers and policy makers
	This study is a forerunner for planned studies incorporating DOW during telephone triage. Therefore, this study gives the foundation for further research, which is why it is suitable as a peer-reviewed publication. We propose that DOW can greatly improve telephone triage, potentially increasing patient safety, however, more research is required.	
	The revised section: <u>Relevance of this study: possible</u> implications for health care providers and policy makers addresses the above.	
	"This study suggests a relation between patients' illness representation, as presented telephonically to an OOH services call handler, and their self-evaluation of urgency, defined as DOW. The relation observed, is that DOW is not random, but follows a pattern, depending on patients' illness representation. This pattern can aid call handlers in understanding patients' perception of urgency, potentially aiding the triage process.	
	This is a new area of research and this study gives direction for future research to further strengthen the evidence. Research on coherence between patient DOW and call handlers', ED and GP physicians' assessment of urgency, both prospectively and	

	retrospectively will strengthen the basis for potential use of DOW as a triage tool. Incorporating DOW as an additional tool in the telephone triage process could potentially aid in the determination of urgency and the type of health care needed, thus increasing patient safety." We believe that the mixed methods approach used, enhances the quality of this study. The thorough deductive analysis of the qualitative data, whereby patterns were identified and themes defined, gives an in-depth insight and understanding of the illness representation of patients to an OOH service, and specifically enhances the quality of this study.	Qualitative data analysis Article summary: Strengths and limitations of this study
In my view, the conclusion is that patients who are insecure about the reason of their problem (week identity, unclear cause, unclear solution) are more worried than patients who do have an idea. I think you should more focus on this in your conclusions. Why is this interesting to know?	The revised section: <u>Relevance of this study: possible</u> <u>implications for health care providers and policy</u> <u>makers</u> (given above) addresses this comment.	Relevance of this study: possible implications for health care providers and policy makers
Before you can say something about safety, you definitely should measure the urgency estimation of medical professionals (you have mentioned this).	The research team is already engaged in further research regarding the incorporation of DOW into telephone triage at OOH services. A prospective cohort study including more than 12,000 callers showed that a high DOW increased the odds for hospitalisation by six-fold (not yet published). In an embedded RCT, no difference in triage response was detected. A process evaluation revealed that this was due to an inadequate intervention. This study, however, further confirms the understanding that patients are able to appropriately estimate the urgency of their health threat. Future research is being planned regarding the examination of difference in triage response with and without DOW, including a new RCT. We hope these future studies will strengthen the basis for the potential use of DOW as a triage tool.	Relevance of this study: possible implications for health care providers and policy makers

methods of this study are weak. Finding relations indicates testing, but this has notsi a a w w this happened. Replacing the word 'interrelation' with 'relation' does not help.	The word "relation" in this manuscript is not used as a statistical term, but in the common literary term according to the Oxford dictionary's definition: "the way in which two or more things are connected; relevance to another." The section: <u>Strengths and imitations of this study</u> states that: " <i>Due to the limited size of the study population, there is a lack of statistical power; however, the results show clear trends and relations, which give direction for future research to strengthen evidence in this new area.</i> " Thus, it is clearly stated that the term "relation" is not	Title Abstract Throughout the paper
w th T w d	o be understood as a statistical term. n analysing the qualitative data, a deductive analysis was conducted, whereby, patterns were identified and hemes defined, as described by Braun and Clarke. <sup>liji</sup> The extent of the analysis and subsequent discussion with a second researcher, using 50% of the study data, resulted in an in-depth insight and understanding of the illness representation of patients to an OOH service.	Strengths and limitations of this study
analysis seems to be poor.	By contrast, the mixed methods analysis comprised grouping of all the calls according to the above defined hemes in each of the five CSM components and isting them according to DOW.	Qualitative data analysis Article summary: Strengths and limitations of this study Mixed methods analysis

Introduction		
I think the information about what you did in this study (in this study, we measured,) fits better in the methods.	In response to this comment, the section: Introduction has been revised. "Telephone triage within out-of-hours (OOH) service is recognised as a mean to reduce pressure and overcrowding of emergency departments (ED) and OOH clinics. <sup>fiv</sup> It aims to assess the urgency of a patient's medical condition in order to determine the correct type of health care needed, thus ensuring patient safety. However, due to the lack of non-verbal cues in telephone consultations, assessing urgency is more challenging than face-to-face consultations. <sup>fv</sup> Studies show that the quality of telephone triage improves with communication between patient and health professional being patient-centred rather than disease-centred <sup>fvi</sup> and that non-normative symptom description and poor communication contribute to under-triage. <sup>fviii</sup> Triage tools, e.g. computerised decision support systems are used to aid the triage process <sup>fviii</sup> ; however, these tools focus on medical information and less on psychosocial or affective information. <sup>lix</sup>	Introduction
	Patients' perception of urgency has previously been examined, comparing ED physicians' and the patients' assessment of the severity of symptoms. <sup>IX Mi</sup> These studies found that patients' perception of urgency can be used as a rough guide to predict the need for hospitalisation. <sup>Ixii</sup> Furthermore, it has been suggested that patients expressing a potential need for hospitalisation should be thoroughly examined for possible severe illness. <sup>Ixiii</sup> Previous studies have also shown that patients' anxiety or worry about a health threat is a major factor in urgent care decision- making <sup>Ixiv Ixv</sup> and that worry is the most important motive for patients contacting OOH services. <sup>Ixvi</sup> Therefore, the measure of a patient's worry about an acute health threat reflects the patient's self-evaluation of urgency. A self-reported verbal ten-point numerical rating scale (NRS) measuring anxiety in patients (1=minimal anxiety to 10=maximal anxiety) has previously been used in several studies in acute care settings. <sup>Ixviii</sup> The anxiety observed in these patients was regarded as acute in relation to the immediate health threat and not due to an underlying psychiatric disease, thus, the feeling of anxiety in this setting was synonymous to worry. <sup>Ixviii</sup> This scale has not been validated. However, as anxiety is a subjective symptom, a subjective scoring system was deemed	

	acceptable. A previous study showed that callers to OOH services were able to rate their degree-of-worry (DOW), using a verbal 10-point NRS (1=minimal worry to 10=maximal worry) as a measure of their self- evaluation of urgency. It was also shown that the DOW scale is feasible for use in large-scale studies. <sup>bixe</sup>	
Results		
I cannot conclude from figure 2 that there is a relation between an illness duration of less than five hours and a low DOW.	In response to this comment, the section: <u>Mixed</u> <u>methods findings</u> have been revised. "A clear trend was observed. Study callers with a medium identity seemed to have a higher DOW, whereas, callers with a strong identity seemed to have a lower DOW and callers with a weak identity generally seemed to have a moderate DOW. There were more callers with a low DOW who had an illness lasting less than 24 hours than callers who had an illness lasting more than 24 hours. Callers with a clear cause for their illness and a clear solution for cure/control seemed to have a low DOW and finally, callers who mentioned a high consequence to their illness seemed to have a high DOW."	Mixed methods findings
Besides, I do not think it is a main finding that "all callers referred to timeline". The call handler asked them.	We consider the fact that all callers referred to the timeline as a main finding, irrespective of the fact that call handlers were instructed to ask callers the duration of their symptoms. In principle, callers could have not answered or not have known.	Quantitative findings according to the CSM framework
Since you did not test the relation between the five components and self- evaluated DOW, I think you should be more careful in the formulation of your conclusions. I would prefer something to say like: it	In response to this comment, the sections: <u>Mixed</u> <u>methods findings</u> (given above) and <u>Summary of main</u> <u>findings</u> have been revised. "Using the five components of the CSM framework, as described by Leventhal <sup>fxx</sup> , our analysis demonstrated that callers presenting their illness to OOH services to a large extent referred to all five components, regardless of their self-evaluated DOW. All callers referred to identity and timeline and callers were least likely to refer to consequence.	Mixed methods findings Summary of main findings

seems that,. (a low DOW was more present,)	Lower DOW seemed to be more present in the group of callers who had a strong illness identity, illness duration of less than 24 hours, a clear cause and a clear solution. Callers who presented a medium or weak illness identity, illness duration of more than 24 hours, an unclear or no cause, unclear or no solution and a perception of high consequence seemed to present a higher DOW." As stated in the section: <u>Relevance of this study:</u> <u>possible implications for health care providers and</u> <u>policy makers;</u> These findings are relevant, because they show a relation between callers' illness representation and their DOW. The relation observed, is that DOW is not random, but follows a pattern, depending on callers' illness representation. This pattern can aid call handlers in understanding callers' perception of urgency, potentially aiding the triage process.	
Discussion		
I still think you should compare your results more with other literature.	In response to this comment, the main author conducted a new literature search on PubMed, using the key words: <i>triage, worry</i> and <i>OOH services</i> . Relevant publications found have been added to the revised section: <u>Introduction</u> (given above) and referenced. This is a new field of research and thereby, existing literature is limited. Patients' DOW in OOH services and EDs has previously been examined with respect	Introduction
	to patient satisfaction, sense of pain, use of OOH services and EDs and urgent care decision-making. To our knowledge, no studies incorporating patients' DOW into the triage process, have previously been conducted.	

Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology. 2006;3:77–101. doi:10.1191/1478088706qp063oa

Würgler MW, Navne LE. Når sygeplejersker visiterer i lægevagten. Copenhagen: Dansk Sundhedsinstitut; 2010.

Leprohon J, Patel VL. Decision-making Strategies for Telephone Triage in Emergency Medical Services. Medical Decision Making. 1995;15:240–53. doi:10.1177/0272989x9501500307

Derkx HP. For your ears only. Quality of telephone triage at out-of-hours centres in the Netherlands. Maastricht: Department of General Practice, University of Maastricht; 2008.

Gamst-Jensen H, Lippert FK, Egerod I. Under-triage in telephone consultation is related to nonnormative symptom description and interpersonal communication: a mixed methods study. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine. 2017;25:52. doi:10.1186/s13049-017-0390-0

Gravsersen DS. Out-of.hours telephone triage by nurses and doctors in Danish acute care settings – A study of quality focusing on communication, safety and efficiency. Aarhus: Research Unit for General Practice, University of Aarhus; 2016.

Holmström I. Decision aid software programs in telenursing: not used as intended? Experiences of Swedish telenurses. Nursing & Health Sciences. 2007;9:23–8. doi:10.1111/j.1442-2018.2007.00299.x

Gifford MJ, Franaszek JB, Gibson G. Emergency physicians and patients assessments: Urgency of need for medical care. Annals of Emergency Medicine. 1980;9:502–7. doi:10.1016/s0196-0644(80)80187-9

Hunt RC, Dehart KL, Allison E, et al. Patient and physician perception of need for emergency medical care: A prospective and retrospective analysis. The American Journal of Emergency Medicine. 1996;14:635–9. doi:10.1016/s0735-6757(96)90077-7

Caterino JM, Holliman C, Kunselman AR. Underestimation of case severity by emergency department patients: Implications for managed care. The American Journal of Emergency Medicine. 2000;18:254–6. doi:10.1016/s0735-6757(00)90115-3

Miyamichi R, Mayumi T, Asaoka M, et al. Evaluating patient self-assessment of health as a predictor of hospital admission in emergency practice: a diagnostic validity study. Emergency Medicine Journal. 2011;29:570–5. doi:10.1136/emj.2010.105247

Booker MJ, Simmonds RL, Purdy S. Patients who call emergency ambulances for primary care problems: a qualitative study of the decision-making process. Emergency Medicine Journal. 2013;31:448–52. doi:10.1136/emermed-2012-202124

Agarwal S, Banerjee J, Baker R, et al. Potentially avoidable emergency department attendance: interview study of patients reasons for attendance. Emergency Medicine Journal 2011;29. doi:10.1136/emermed-2011-200585

Keizer E, Smits M, Peters Y, et al. Contacts with out-of-hours primary care for nonurgent problems: patients' beliefs or deficiencies in healthcare? BMC Family Practice 2015;16. doi:10.1186/s12875-015-0376-9

Fosnocht D, Swanson E. Pain and anxiety in the emergency department. Annals of Emergency Medicine. 2004;44. doi:10.1016/j.annemergmed.2004.07.284

Craven P, Cinar O, Madsen T. Patient anxiety may influence the efficacy of ED pain management. The American Journal of Emergency Medicine. 2013;31:313–8. doi:10.1016/j.ajem.2012.08.009

Gamst-Jensen H, Huibers L, Pedersen K, et al. Self-rated worry in acute care telephone triage: a mixedmethods study. British Journal of General Practice Published Online First: December 2018. doi:10.3399/bjgp18x695021

Leventhal H, Phillips LA, Burns E. The Common-Sense Model of Self-Regulation (CSM): a dynamic framework for understanding illness self-management. Journal of Behavioral Medicine. 2016;39:935–46. doi:10.1007/s10865-016-9782-2

#### **VERSION 3 – REVIEW**

REVIEWER	Ellen Keizer
	Institute for Primary Care, Zürich University, Switzerland
REVIEW RETURNED	21-Jun-2018
GENERAL COMMENTS	I thank the authors for responding to my comments. I am satisfied
	with the answers and recommend acceptation!