An insight to nurse workload: predicting activities in the next shift and analyzing bedside alarms influence

Renata M. de Carvalho, Huyen Nguyen, Maikel Heetveld
Department of Mathematics and Computer Science
Eindhoven University of Technology
r.carvalho@tue.nl, {k.h.nguyen, m.m.a.r.heetveld}@student.tue.nl

Jolanda Luime
Máxima Medisch Centrum
jolanda.luime@mmc.nl

Abstract

The effects of nursing shortage and the increasing nursing workload have been emphasized in recent years. Hospitals and healthcare organizations have actively implemented initiatives to combat the workforce shortage and mitigate the effects of high workload among their nurses. In cooperation with a Dutch hospital, this project seeks to understand the workload of nurses by understanding what will happen in the next shift and by measuring the alarm workload of their nurses, particularly the bedside alarms that send signals to wifi phones of nurses, to obtain a partial workload of nurses that may reflect the intensity level of the total nurse workload. The project also investigates the effects of the alarm workload, patient, working condition and staff individuality on nurse performance. Finally, the project seeks to discover possible contributing factors to the alarm occurrence, therefore finding a mitigation to reduce the bedside alarm workload.

1. Introduction

Although nurses are the largest section of the health profession [1], hospitals and care providers in the Netherlands are under pressure of the nursing shortage [2]. In 2019, the healthcare sector recorded 39.000 unfilled vacancies which rose by 37% compared to demand in 2018 [2].

The nursing shortage leads to low-quality patient service. Under the nursing shortage, nurse-to-patient ratio is high [1]. In order to maintain the nurse-to-patient ratio at an acceptable level, the hospital is forced to close beds. So, rather than increasing the patient numbers, beds are closed to prevent work overload. However, within the current system still workload can raise as complicated patients are together on the ward. This could be because 2 nurses needs to attend one patient, but also by complex wounds, deteriorating of vital signs, wandering patients. Studies have shown that the nursing shortage could cause errors, high mortality and morbidity rates to patients [1, 3].

Not only does the nursing shortage affect patients, it also causes nurses to experience increased workload, burnout and job dissatisfaction [1]. More recently, the hospital has also requested nurses to come to work on a free day. In a 3 month period in 2019, the hospital did that 258 times. Still nurse-patient ratio is satisfied but the nurse is forced to change his/her private schedule. Those effects on nurses could eventually lead to profession leaving of nurses, thus creating even bigger nursing shortage. A survey by NCBI [4] reported that the Netherlands had almost 30% nurses working overtime, 55% nurses evaluating their jobs as busy and 10% nurses considering their jobs as too busy. Another survey by Schaufeli [5] reported that among 4.000 healthcare workers, 15% of them were exposed to the risk of burnout while 7% of them were already on sick leave due to symptoms of burnout. With low job satisfaction, nurses tend to leave their jobs [1]. 10.4% nurses in the Netherlands reported their intention to leave the profession [5].

The goal of this project is tow-fold: 1) to understand the workload faced by nurses, and 2) to be able to predict what is expected in the next shift. In this context, actions can be taken to not overload nurses. In order to address the first point, we will investigate the alarm workload of nurses. Particularly the bedside alarms that send signals to wifi phones of nurses, in order to obtain a partial workload of nurses that may reflect the intensity level of the total nurse workload. The second point will be addressed by creating a first prediction model able to predict the amount of tasks that is expected in the next shift.

With the bedside alarms investigation, we aim at identifying which characteristics of patient-care related workload potentially increase the perceived workload of nurses. Bedside alarms are alarms generated by patients through an alarm button located on the side of their beds. Patients ring off alarms when they need help from nurses. This might happen when the patient is in a life-threatening situation or when the

HĭCSS

patient needs help to go to the toilet. Furthermore, identifying the characteristics that would increase the perceive workload may allow for a better distribution of patients that fall into the category of the identified characteristics among the different nurses of the same shift.

The alarm is not the only factor that reflects the workload of nurses. For instance, it was pointed out by the interviews conducted in this study that discharging a patient (and all the reports and other activities involved) is something that takes time of a nurse and ideally it would be nicely distributed among nurses. In this context, as an ideal solution (aimed at future stages of this project), it is possible to predict the amount of each type of activity that will happen in the next shift. In this paper, we will also present an initial model that is able to predict the amount of activities in general. This is already a good step towards the solution, as it is able to capture many interesting patterns.

The remainder of the paper is organized as follows. Section 2 gives an overview of existing initiatives to solve the nursing workload problem. Section 3 explains the data we had access to, while Section 4 details how we handle this data to both predict the activities of the next shift and to identify influencing factors related to bedside alarms that might influence the way nurses perceive the workload. Finally, Section 5 concluded the paper with a summary of our findings.

2. Literature Review

Many studies have been undertaken to define and measure the nursing workload. Medical Subject Heading Terms database (MeSH database) of the US National Library of Medicine introduced in 1982 the term workload as "the total amount of work to be performed by an individual, a department, or other group of workers in a period of time". In 1997, the work of Needham [6] defined the nursing workload as "the total nursing time consumed on direct, indirect patient care, and other non-patient related activities, such as ward and organisation management".

Concerning how to quantify the workload of nurses, the Therapeutic Intervention Scoring System - 28 (TISS-28) [7] is a commonly used standard in the Netherlands. In the TISS-28 system, which is an evolution of the Therapeutic Intervention Scoring System (TISS) [8], all tasks to be done by a nurse should be graded a specific number of points (0 to 3 on multiple aspects). In the case of TISS-28, this is a total of 28 different tasks, which gives a total of 88 points; other tasks need to be ordered under the most representative available. Each point is equal to 10.6 minutes of nursing time spent on direct patient care. Hence, the points of these tasks should together add up to 46, which corresponds to a full eight-hour shift. This method has two major downsides. First, it is specifically tailored for the ICU, which is not comparable to other departments. Second, this whole methodology depends on the documentation of nurses, requiring them to perform an additional task (after their shift).

Alternatively, there is the RAFAELA system [9]. This is a standard for the ranking of workload for nurses. The RAFAELA system measures nursing intensity by ranking six areas that require nursing intervention of patient needs. Each of these areas is ranked from A=1 to D=4, with A being patients that manage relatively well on their own and D, patients cannot manage without care. This results in a score varying from 6 to 24 per patient for the last 24 hours. However, the RAFAELA system again requires nurses to document their shift. This documentation is currently not being done in our study, as such we had no data available about this documentation.

In terms of dimensions that contribute to nursing workload, several of them have been proposed by Crayon and Alvarado [10] including physical, cognitive, time pressure, emotional, quantitative, qualitative, and variability dimensions. This work is reviewed by Cordova *et. al.* [11], which summarized the physical workload as manual direct tasks such as moving and lifting patients and physical organization associated with patient care such as gathering intravenous pumps and vital signs monitoring equipment. Cordova *et. al.* also stated that "the cognitive workload is the measurement of intellectual processing of patient information that drives performance and decision; the time pressure is the demand of number of tasks performed under temporal constraints; the emotional workload is the demand of dealing with emotional issues, such as patient death, end-of-life care, and family demands; the quantitative and qualitative workload are the amount and difficulty of work, respectively; the variability workload is defined as the fluctuation of workload during a time period".

As a study of the impacts of nursing workload, particularly of alarm-related workload to nurse performance and nurse experience, Shanmugham *et. al.* [12] established experiments of modifying the default alarm settings and evaluated the perceived workload of nurses under different settings. This study also assessed the relationship between perceived workload, response rate and caregiver experience.

Alarm fatigue, its contributing factors and its influence on staff performance, have been also frequent subjects of study. Sendelbach and Funk [13] have

demonstrated that 72% to 99% of clinical alarms are false alarms, leading to alarm fatigue. Alarm fatigue is sensory overload when clinicians are exposed to an excessive number of alarms, which can result in desensitization to alarms and missed alarms. The work of Deb and Claudio [14] proposed a model studying the significance of working conditions and staff individuality on alarm fatigue and, consequently, alarm fatigue on staff performance. Results show that staff performance deterioration is influenced by a combination of alarm fatigue, working conditions and staff individuality. Nurses and their response time have shown to not be influenced by alarm fatigue. Finally, working conditions and staff individuality have shown to contribute to alarm fatigue [14].

In this research, we not only analyze the conditions of nurses and the relations such conditions might have with nurse workload. Rather, here we focus on predicting what will happen in the next shift, focusing on the amount of tasks expected in the next shift. In addition, we do a first attempt on identifying which factors might influence the amount of bed side alarms. As a long term goal, we aim at predicting the amount of the different types of tasks in the next shift, as we know from previous works that they have different influences and should be distributed among nurses differently.

3. Data Description

We had access and analyzed six datasets, called: *meta*, *form*, *door*, *wifi*, *alarm* and *patient admission*. The content of each dataset is described in the following. Also, it is noted that the datasets belong to only general clinical or inpatient departments. It excludes ER and ICU, day care and short stay, as these units are different from the general departments in their procedures and bedside alarm regulation.

1. ***Meta* dataset** During each working shift, nurses are required to perform administration tasks over the profiles of patients. The administration includes: making activity plan; checking admission, medicines and allergies of patients; writing reports about problems and progress of patients; and many more. Those activities of nurses are performed via various modules of an internal patient system of the hospital.

   The *meta* dataset records administrative activities of nurses, including nurse identifier, timestamp of the activity, the specific module name describing the activity, and patient identifier. It does not contain details on the activities itself, but relates to views and actions in the electronic patient system.

Note that all personal data (e.g., nurse identifier, patient identifier) were anonymized.

2. ***Form* dataset** *Form* dataset is similar to the *meta* dataset. While the *meta* dataset describes the used modules in the patient system, the *form* dataset shows which specific forms within those modules are administered by nurses. The forms could be about: registering patient problems; patient background information such as alcohol and smoking history; and patient visit notes.

   The *form* dataset records form-related activities of nurses including the nurse id, timestamp of the form activity, the form name for the specific activity type, and the patient id that the form belongs to.

3. ***Door* dataset** Nurses must use their personal keys in the form of identifier cards to access restricted areas within the hospital such as medicine rooms, doctor wards, and personnel rooms. Each time using the key to open doors of those areas, nurses are registered in the door system about their use.

   The *door* dataset records the door use of nurses including the nurse id, timestamp of opening the door and the door location.

4. ***Alarm* dataset** Many medical devices within the hospital are built with alarm functions that create alarms in audible, visual and other forms to communicate to nurses alarming situations and deteriorating conditions of patients (e.g., a life-threatening arrhythmia, or the telemetry system's low battery condition). Those devices can include infusion pumps, respiratory monitoring equipment, feeding pumps, cardiac monitors, ventilators, bed and chairs. This paper focuses on analyzing alarms generated from patient beds. Patients can press an alarm button located on the side of their beds to signify nurses when they need help in activities such as walking, toilet, bathroom, bed-chair transfer, thus triggering bedside alarms.

   The *alarm* dataset records the patient id, timestamp when the alarm was generated, and the location where the alarm is generated (department, room and bed number).

5. ***Patient Admission* dataset** When patients are admitted to the hospital, they are registered in the hospital system with their information. For this paper, only basic information of the patient are provided.

The *patient admission* dataset records patient id, patient age, patient gender, patient body mass index (BMI), reason of the patient admission, start and end date of the patient admission, location, room and bed of the patient, and start and end date of any specific treatment on the patient during this admission.

4. Analyzing nursing workload

As previously mentioned, we aim at analyzing "low level" events performed by nurses, without asking them to report what they are doing, in order to understand the events that are taking place (see Section 4.1) and to be able to get insights about what will happen in the next shift. In this way, the hospital should be able to better distribute the workload among nurses. We will approach this problem from two different angles: 1) from the amount of activities a nurse performs in one shift (Section 4.2); and 2) from the burden that alarms raised by patients (called bedside alarms) can generate, specially when alarms interrupt the activities being performed (Section 4.3).

4.1. From low level events to high level activities

We identified seven main digital activities in the hospital that can be performed by nurses, namely *admission*, *daily process*, *medication*, *patient status*, *patient orientation*, *discharge* and *general work*. They are considered here as *high level* activities. Please note that these are activities that can be identified through the use of the system. We stress that there are still "invisible" activities performed by nurses that are not possible to be identified as there is no register in the system that they are performed, e.g., helping a patient to use the toilet.

In order to identify which of these activities is being performed, we first defined, for each *low level* event that we can find in the dataset, to which activities they might be linked to. For instance, a low level event might only indicate that the nurse filled in some information about the smoking conditions of the patient. It is possible that a low level event can be performed in the context of more than one activity. Thus, as the high level activity is a set of low level events, the approach to define which activity is being performed is the following:

1. **Group events by five-minute time difference:** For each nurse, on each shift, on each patient, follow the timeline of events performed in the *meta* and *form* datasets. If two neighboring events have less than or equal to five-minute difference

in timestamp, then the two events are grouped together into the same time period. For instance, the *admission* activity should comprise of a set of events that indicate the nurse is filling in many forms about prior and/or living conditions of the patient. When such events happen within 5 minutes, we consider them as the being part of the same *admission* activity. If there is a gap of more than 5 minutes between them, they will be considered as part of different activities.

2. **After grouping events by five-minute difference:** Retrieve the most dominant activity in each group. As said, we identified that some events might be related to more than one activity. For example, if in a given period, *medication* can be linked to 3 out of 5 events, and all other activities can only be linked to less than that, this is the dominant activity and then considered as the activity being performed. As most of the events can only be linked to one activity, it is very likely that there is only one dominant activity.

3. **Duration of the activity**: We consider the timestamp of the first event in the group as the beginning of the activity, and the timestamp of the last event in the group as the end of the activity (although we know this is not true). The timestamp of the last event only indicates when the last event that is part of the high level activity started, it is still unclear how long this event takes.In Section 4.3), we discuss this further and we explain how we address this issue.

4.2. Amount of activities performed by a nurse

As a first step, we developed a simple prediction model using Random Forests regression, which is a supervised learning algorithm that uses ensemble learning method for regression. As Random Forest is a decision tree-based technique, it allows us to investigate the final tree and use it as knowledge to identify relevant characteristics to predict how busy the next workload will be.

As features, we used information about the previous shift, e.g., number of patients, average age of patients, number of patients that need intravenous therapy, number of patients that need medicine, total number of "high level" activities performed (see previous section). Our target prediction is the number of activities that will be performed on the next shift. We believe that this would give the hospital a better idea about the number of nurses needed in the next shift, in a way that

nurses could handle tasks without a feeling that they are overloaded.

Figure 1 shows that this prediction is possible and the model is already able to identify seasonality related to the shifts themselves. Morning shifts have a higher amount of tasks, while night shifts are relatively low.
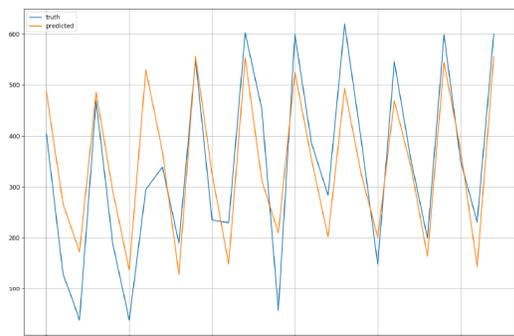

Figure 1: Prediction of number of activities that will take place in the next shift. The y-axis represents the number of activities, while the x-axis represents each shift in the test set, ordered by time.

Also interesting to see in Figure 1 is that some hospital policies are also captured. For instance, it is common for the hospital to plan more elective patients to be admitted on Tuesdays, so that patients would probably be discharged before the weekend. The Random Forest model can also capture that. In the test set shown in the figure, the fifth peak represents a Tuesday morning. One can see that the peaks in the prediction line slightly increase until the fifth peak and slightly decrease after that.

Predicting the number of activities in the next shift is useful, however, having more insights on different aspects of the expected activities might help even more. Having said that, we will investigate further one of the aspects involved to the activities and to the nurse workload: the number of bedside alarms.

4.3.   Bedside alarm workload

In the context of the workload related to bedside alarms, we distinguish two kinds of alarms: *interrupting*, and *non-interrupting* alarms. We considered as *interrupting* alarm, any alarm that falls in one of the following conditions:

- The time a nurse takes to answer the alarm is longer than a threshold. As the nurse cannot assume that the ringing alarm is not urgent, and he/she still takes longer than the threshold to answer it, we consider that what he/she is

currently doing cannot be paused. Thus, we consider that the alarm is interrupting the current work of the nurse. This project defined the median response time as the threshold. Therefore, if an alarm has the response time longer than the median response time of all alarms in the dataset, then the alarm is an interrupting alarm. The median value is used because of the non-normally distributed response time we had.

We calculate the response time as the duration in minutes since the timestamp when the alarm first rang until the timestamp of the last time the same alarm rang, found in each alarm record. Note that the the first and the last alarm timestamp are different than the first and the last alarm signal timestamp. As shown in Figure 2, there is a latency between the first alarm timestamp and the first alarm signal. In case a nurse accepts the alarm just right before the one-minute interval alarm signal, the last alarm signal is the one prior to the acceptance time. There is also a latency between the alarm acceptance timestamp and the last alarm timestamp. The latency happens due to delay in data transmission.

- The time difference between the current procedure the nurse is performing and the alarm is negative. If the time difference is negative, the current procedure is considered as uncompleted. Thus, when the alarm occurs, it is interrupting the work of the nurse. We calculate this time difference by subtracting timestamp of the last low level event of the current activity timestamp from the first alarm timestamp, as illustrated in Figure 3.

The current system of *meta* and *form* datasets can only capture digital events of nurses for administering patients. Nurses can perform other types of activities that are not visible and recorded by the system. Therefore, it is critical to acknowledge that the last event timestamp is not the timestamp for completing the procedure, rather it only indicates when the last event that it part of the activity started. One needs to be aware of the time gap between the last event timestamp of the procedure and the last timestamp of the procedure and this gap is unknown.

In addition, the *door* and *wifi* datasets give us an indication on where (which room) the nurse is at a certain moment in time. If no event is captured in the system that can be related to that location, we can only assume the nurse is performing one of these "invisible" tasks.
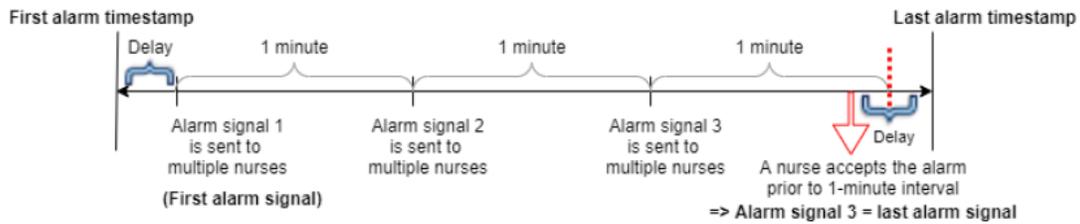
Figure 2: Response time of the alarm: first and last timestamp of the alarm versus first and last signal of the alarm
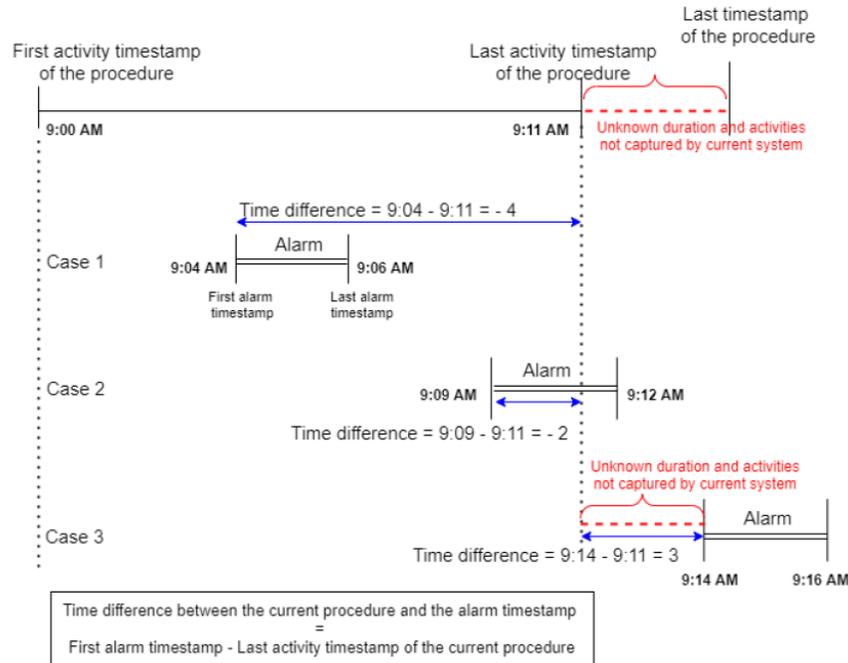

Figure 3: Illustration for retrieving the time difference between the current procedure and the alarm timestamp

- The nurse that responded to the alarm performs the same high level activity on the same patient repeatedly 30 minutes before, during, and after the alarm timestamps. Thus, we check if 30 minutes before, during, and after the alarm occurrence, the nurse responding the alarm has been continuously performing events on the same patient id. Due to the alarm occurrence, the nurse may need to attend the alarm, the work for the current patient is delayed/interrupted until the nurse finishes the alarm handling. Opposite to the definition on the first point, here we assume that the current activity can be paused and it is for the handling of the alarm, later this activity is resumed. The 30-minute range is used as an approximation of the alarm handling duration that the nurse spends before returning to the current patient. Also this 30-minute range is used to find the latest patient id that the nurse is attending. Figure 4 illustrates this scenario.

4.3.1. Factors influencing the workload

In order to identify factors that directly influence the number of alarms triggered and/or the performance of nurses due to the alarms triggered by patients, we performed 53 Linear Regression (LR) tests, which attempts to explain the relationship between two variables using a straight line:

- 20 LR tests between patient characteristics (age, BMI, length of stay, time since admission, length of nurse forms) and alarm measures (number of alarms, number of signals, number of interrupting
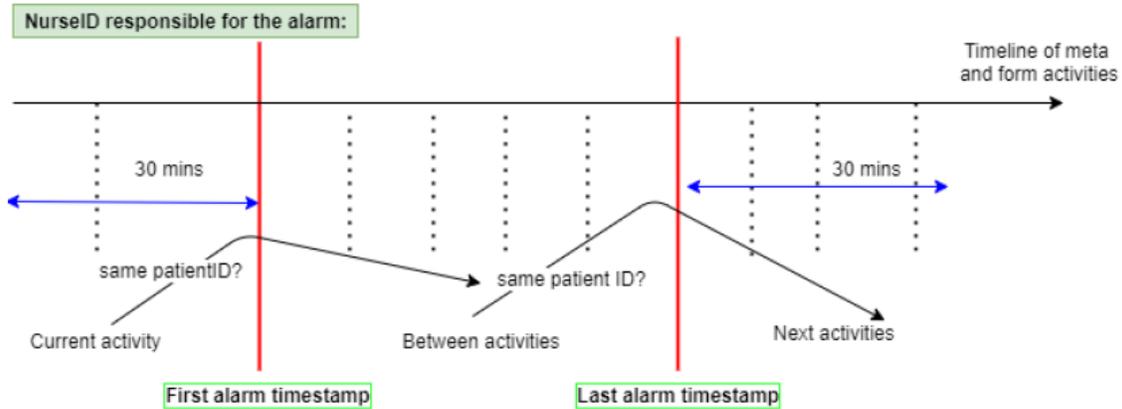
Figure 4: Illustration for retrieving the similar activities before, during and after an alarm.

alarms, number of non-interrupting alarms);

- 12 LR tests between working condition measures (number of patients, number of nurses, nurse-to-patient ratio) and alarm measures;

- 4 LR tests between alarm measures and performance measures (response time);

- 5 LR tests between patient characteristics and performance measures;

- 7 LR tests between measures related to the activities being performed (time difference between the current activity and alarm) and performance measures;

- 5 LR tests between measures related to the location where activities are being performed (distance from the location of the current activity to the location where the alarm was triggered) and performance measures.

It is worth to mention that before applying the Linear Regression tests, we removed outliers based on Z-score and interquartile range (IQR). After removing outliers, we split our dataset into training and test sets, using a ratio of 80:20. Although all 53 tests were performed, in the following, we will present only the most relevant results. The linear regression tests only consider one dimension at a time, we will also investigate multiple dimensions at a future stage.

Number of interrupting alarms and response time
The number of interrupting alarms is hypothesized to have influence on the response time of nurses. The Null hypothesis is that the number of interrupting alarms does not have influence on the response time. A LR

Table 1: Performance metrics from LR test between number of interrupting alarms and response time

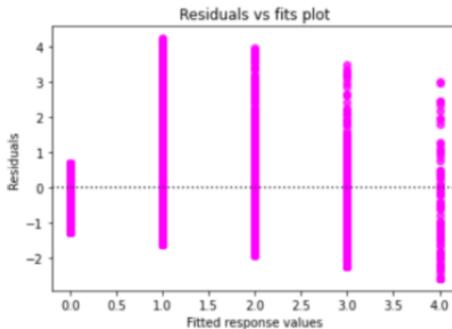| Predictor	Number of interrupting alarms
Response	Response time
Coefficient $\pm$ STD	0.3339 ± 0.027
P-value	0
R2	0.071
MSE	1.728

test is conducted to verify this hypothesis. The test is setup with hour-based calculation that checks how many alarms within an hour that interrupt the work of each nurse id (the number of interrupting alarms that a nurse has within an hour) and what the response time the specific nurse has towards alarms of that hour.

Table 2 presents the result of the SLR test between the number of interrupting alarms and the response time. With the p-value of 0, thus less than 0.05, the Null hypothesis is rejected. The number of interrupting alarms is likely to have effect on the response time.

Although the hypothesis is accepted that the number of interrupting alarms has influence on the response time, the strength of the acceptance is still low. The R-squared value of 0.071 means that only about 7% of the variance in the response time is explained by the number of interrupting alarms. The MSE of 1.728 is relatively high comparing to the general response time range between 0.45 and 2.45 minutes. Assumptions about the linear relationship and the normality of residuals are weakly accepted based on the data scatter plot and the normal probability plot respectively in Figures 5a and 5c. On the other hand, assumptions about the equal variance and independence of errors can be accepted based on the residuals vs fits plot in Figure 5b.
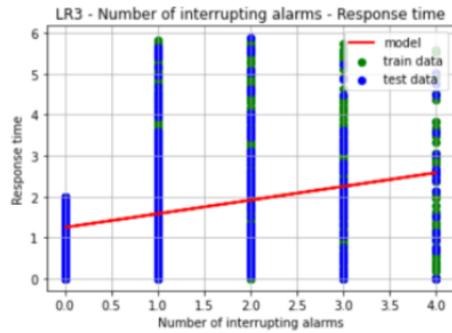
(a) Data distribution

(b) Residuals vs fits
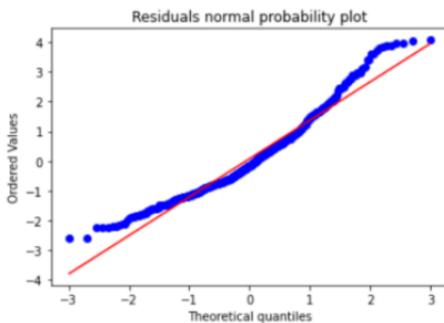
(c) Residual normal probability

Figure 5: Accepted association: number of interrupting alarms and response time.

It is inferred from the LR test result that the more interrupting alarms the nurse has within an hour, the longer the nurse takes to respond to alarms of that hour. For every one unit increase of the number of interrupting alarms, there is an increase of the response time by 0.3339 times (based on the coefficient value). It is sensible to see this result. While multiple alarms are interrupting the nurse work within an hour, the nurse may not be able to response quickly due to either multiple alarms in the hour, or the nurse work being proceeding.

Table 2: Performance metrics from LR test between time difference between the current activity (patient status) and alarm timestamp and response time

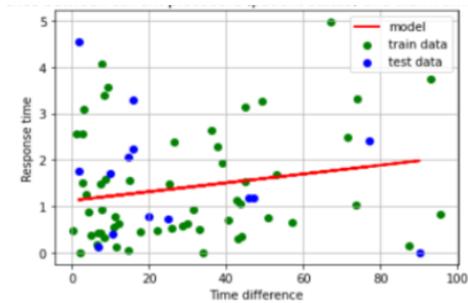| Predictor	Time difference between current activity (patient status) and alarm timestamp
Response	Response time
Coefficient ± STD	0.0095 ± 0.006
P-value	0.142
R2	0.04
MSE	1.776

Time difference between current activity (patient status) and alarm timestamp and response time It is expected that the current activity of a specific nurse affects the alarm response time of that nurse. If the current activity is not completed yet, the nurse may have long response time towards alarms. The time difference between the current activity and the alarm timestamp is the difference in minutes since the last known event of the current activity of the nurse until the first alarm timestamp.

Here we present the result of the LR test that hypothesizes that the time difference between the current procedure of patient status and the alarm timestamp have influence on the response time. The Null hypothesis states that the time difference between the patient status procedure and the alarm timestamp does not have influence on the response time. Table 2 describes the result of the LR test. With the p-value of 0.142, the Null hypothesis can be accepted. However, after verifying SLR assumptions, the association could still be accepted weakly that the input measure has influence on the response time. Figure6 shows an indication of the linear relationship, the equal variance and independence of errors.

It is inferred from the LR test result that the longer the time difference between the current activity of patient status and the alarm timestamp, the longer the response time. For every one unit increase of the time difference, there is an increase of the response time by 0.0095 times (based on the coefficient value). This seems to be less intuitive because if the time difference is long, the current activity (patient status) should have been completed already, thus the response time should be short.
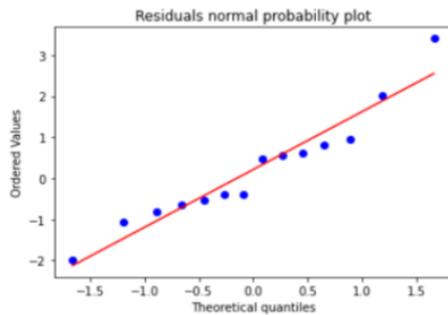
5. Conclusions

This paper seeks to evaluate the concern of nurses about the high workload during their hospital shifts

(a) Data distribution
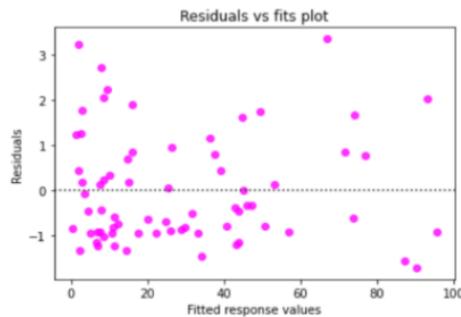
(b) Residuals vs fits

(c) Residual normal probability

Figure 6: Accepted association: Time difference between current activity (patient status) and alarm timestamp.

and help the hospital in finding a mitigation to reduce the workload. Based on the literature review and the available data of the hospital, this project is scoped to measure the workload via patient-related care, particularly, by predicting the amount of activities expected for the next shift and by analyzing the bedside alarms that send signals to wifi phones of nurses to retrieve a partial workload of nurses as a way to reflect the total nurse workload.

We showed that we are able to predict the amount of activities a shift will have based on the information about the previous shift, after a prediction model is trained. When analyzing the factors that influence the performance of nurses when dealing with multiple alarms, that sometimes interrupt the current work being performed by nurses, we were able to find associations between some of the factors analyzed, as presented in Section 4. It is noteworthy that the discovered associations still show weak indication. The results of the analysis should be used as the initial groundwork for future research of the hospital.

The complexity but also the limitation of the data system have posed a lot of work to improve the research in the future. The alarm workload is currently measured under the quantitative value. However, it should also be measured under the qualitative values based on the time, physical and mental requirements. Specially because the main burden nurses suffer is the mental workload created by the noise created by the alarms and in a very high frequency. Although in this paper we defined a concept of interrupting alarm, it is worth making an effort to see the problem from the perspective of a nurse: 1) there is a situation being handled already by the nurse regarding a patient; 2) the phone in the nurse's pocket is constantly ringing, producing a loud and disturbing noise in the room; and 3) the nurse now needs to concentrate on the current task and deal with the fact that there is another patient with a potential emergency situation. With this, the nurse must now reason over the whole situation and decide what action must be taken next: stay with the current patient or move to the one calling.

A field research will be conducted by observing, interviewing and surveying the nurses during their working shifts. In addition, we will improve our method to incorporate more information about the location of nurses, as this might add a lot of information concerning which activities are indeed being performed, and might also allow us to identify some of the "invisible" tasks we mentioned before.

References

[1] L. M. Haddad, P. Annamaraju, and T. J. Toney-Butler, "Nursing shortage," *Treasure Island (FL): StatPearls Publishing*, 2020.

[2] J. Visser, "The netherlands has an urgent need for cops, nurses and teachers." The Holland Times, 2019. https://www.hollandtimes.nl/articles/national/the-netherlands-has-an-urgent-need-for-cops-nurses-and-teachers/.

[3] W. Moloney, D. Gorman, M. Parsons, and G. Cheung, "How to keep registered nurses working in new zealand even as economic conditions improve," *Human resources for health*, vol. 16, p. 45, September 2018.

[4] W. H. O. R. O. for Europe, E. O. on Health Systems, Policies, A. M. Rafferty, R. Busse, B. Zander-Jentsch, W. Sermeus, and L. Bruyneel, *Strengthening health*

systems through nursing: evidence from 14 European countries. Health Policy Series: 52, World Health Organization. Regional Office for Europe, 2019.

[5] W. B. Schaufeli and D. V. Dierendonck, "A cautionary note about the cross-national and clinical validity of cut-off points for the maslach burnout inventory," *Psychological Reports*, vol. 76, no. 3_suppl, pp. 1083–1090, 1995.

[6] J. Needham, "Accuracy in workload measurement: a fact or fallacy?," *Journal of Nursing Management*, vol. 5, pp. 83–87, 1997.

[7] D. R. Miranda, A. de Rijk, and W. Schaufeli, "Simplified therapeutic intervention scoring system: The tiss-28 items–results from a multicenter study," *Critical Care Medicine*, vol. 24, pp. 64–73, 1996.

[8] D. J. Cullen, J. Civetta, B. A. Briggs, and L. Ferrara, "Therapeutic intervention scoring system: a method for quantitative comparison of patient care," *Critical Care Medicine*, vol. 2, pp. 57–60, 1974.

[9] M. Andersen, K. Lønning, G. M. W. Bjørnelv, and L. Fagerström, "Nursing intensity and costs of nurse staffing demonstrated by the rafaela system: liver vs. kidney transplant recipients.," *Journal of Nursing Management*, vol. 24, pp. 798–805, 2016.

[10] P. Carayon and C. J. Alvarado, "Workload and patient safety among critical care nurses," *Critical Care Nursing Clinics of North America*, vol. 19, no. 2, pp. 121–129, 2007. Safe Patient Handling.

[11] P. B. de Cordova, R. J. Lucero, S. Hyun, P. Quinlan, K. Price, and P. W. Stone, "Using the nursing interventions classification as a potential measure of nurse workload," *Journal of Nursing Care Quality*, vol. 25, no. 1, pp. 39–45, 2010.

[12] M. Shanmugham, L. Strawderman, K. Babski-Reeves, and L. Bian, "Alarm-related workload in default and modified alarm settings and the relationship between alarm workload, alarm response rate, and care provider experience: Quantification and comparison study," *JMIR Hum Factors*, vol. 5, p. e11704, Oct 2018.

[13] S. Sendelbach and M. Funk, "Alarm fatigue: A patient safety concern," *AACN advanced critical care*, vol. 24, pp. 378–86, 10 2013.

[14] S. Deb and D. Claudio, "Alarm fatigue and its influence on staff performance," *IIE Transactions on Healthcare Systems Engineering*, vol. 5, pp. 183–196, 07 2015.