UNIVERSITY OF TWENTE.

**Faculty of Behavioural, Management and Social Sciences**

# Detection of deterioration in continuously monitored surgical ward patients

## Analysis of a new risk prediction model combining nurse intuition and physiological changes

**Marlous Kalsbeek**

**MSc Thesis**

**Health Sciences – Personalized monitoring and coaching**

**September 2020**

**Supervisors University of Twente:**
Dr. M. (Marleen) Groenier
M.C. (Mathilde) Van Rossum

Master Thesis Health Sciences

# Detection of deterioration in continuously monitored surgical ward patients

Analysis of a new risk prediction model combining nurse intuition and physiological changes


| | |
|---|---|
| **Author** | Marlous Kalsbeek |
| **Student number** | S2177935 |
| **Master** | Health Sciences<br>Track: Personalized monitoring and coaching |
| **Institution** | University of Twente<br>Faculty of Behavioural, Management and Social Sciences |
| **Supervisors**<br>**University of Twente** | Dr. M. (Marleen) Groenier<br>M.C (Mathilde) Van Rossum |



UNIVERSITY OF TWENTE.

# ABSTRACT

## BACKGROUND

Due to the ageing population, the demand for intensive care is increasing (1). As a result, critically ill patients will be transferred to the general surgical ward earlier. These critically ill patients are often complex cases and early detection of deterioration is critical in preventing major complications that can result in impairment, function loss, longer days spend in the hospital, higher hospital costs and even mortality. Nurses monitor patients by manually measuring vital signs. Literature shows that changes in vital signs can detect deterioration of a patient. Nurses use an instrument called MEWS to score deviating vital signs and detect deterioration. However, it turns out that nurse intuition plays an essential role for nurses in the identification of deterioration and that vital signs are mostly used to validate their intuitive feeling. Literature even shows that nurse intuition alone is better at predicting deterioration than deviating vital signs. Unfortunately the current method of measuring and recording vital signs is suboptimal and importance of nurse intuition is often not acknowledged sufficiently, resulting in an underestimation of physiological deterioration in patients. Continuous monitoring of vital functions can potentially resolve the shortcomings of the current practice and has already shown promising results in the early detection of deterioration. Though, it has not yet been established whether combining nurse intuition and continuous monitoring will yield even more benefits.

## OBJECTIVE

The research objective is firstly, to investigate the predictive value of MEWS and nurse worry indictors in detecting patient deterioration. Secondly, it will be investigated whether continuous vital sign monitoring combined with nurse worry results in earlier and better detection of deterioration in surgical patients compared to current practice.

## METHOD

An exploratory study was conducted. In the first part, the current performance of the MEWS and nurse worry indicators as composite scores were investigated and compared by calculating sensitivity, specificity, positive- and negative predictive values and Area under the Curve values on a discretely measured dataset. In the second part, logistic regression was used to determine the predictive performance of the separate variables in predicting adverse events. The variables with highest individual predictive values were combined in a multi logistic regression model to construct a new risk scoring instrument that was then tested on continuously recorded vital signs data of the patients.

## RESULTS

Results show that MEWS and nurse worry as composite scores have poor predictive power with an AUC of 0.6 and 0.5, respectively. The most important variables that make up MEWS and Nurse worry, were blood pressure (P=0.006) , heart rate (P≤0.000), respiratory rate (P=0.095), nurse worry about temperature (P=0.168), nurse feeling (P=0.270) and patient feeling (P=0.169). These variables were used in the new model (CWS). The CWS with a time interval of 60 minutes had a sensitivity of 77% , specificity of 72%, PV+ 44 % , a PV- of 92%, and an AUC of 0.77. These results are significantly better than the results observed for traditional MEWS on both manually obtained data and on the continuous data. The CWS detected adverse events on average a day earlier than the traditional MEWS on manually obtained data.

## CONCLUSION

Introducing continuous monitoring on the surgical wards and using vital signs and nurse worry indicators to detect deterioration in patients' condition both show great potential towards earlier and better detection of adverse events. Better and earlier detection of an adverse event can results in more patient safety, lower mortality rates, reduce workload of nurses and improve communication between health professionals. Future research should focus on establishing more certainty about scoring ranges for the vital signs in continuous monitoring solutions, automated pattern recognition in vital signs in combination with nurse worry indicators and should look into the added value of relative scores besides absolute scores to indicate deterioration.

# PREFACE

This thesis is written for the master Health Sciences at the University of Twente. The thesis assignment was part of a collaboration of ZGT Hospital and researchers of the University of Twente.

At the beginning of writing this thesis no one could have predicted how different the world would be by the time I ended my thesis. There were many changes and setbacks during my thesis writing period. Not only was there a global pandemic that changed our way of life, but there were also some personal loses which had great impact. Luckily, I had some great people around me that helped me to finally finish my master thesis.

First of all, I would like to thank my supervisors from the university of Twente, Mathilde van Rossum and Marleen Groenier. It was a very strange time for us all but we all managed the make it work. I'm especially thankful for their time, patience and clear perspectives.

Secondly, I would also like to thank my friends and family that have helped me stay positive, discusses ideas with and take my mind off when needed. I feel very blessed to have them in my life.

Marlous Kalsbeek

Enschede, September 2020

# ABBREVIATIONS

| | |
|---|---|
| AUC | Area Under the Curve |
| CCO | Critical Care Outreach |
| CMEWS | Continuous Modified Early Warning Score |
| CWS | Continuous Warning Score |
| DENWIS | Dutch Early Nurse Worry Indicators Score |
| EPV | Events Per Variable |
| EWS | Early Warning Score |
| FN | False Negative |
| FP | False Positive |
| HR | Heart Rate |
| ICU | Intensive Care Unit |
| MET | Medical Emergency Team |
| MEWS | Modified Early Warning Score |
| NW | Nurse Worry |
| PV- | Negative Predictive Value |
| PV+ | Positive Predictive value |
| RR | Respiratory Rate |
| RRS | Rapid Response System |
| RRT | Rapid Response Team |
| SPO2 | Saturation |
| TP | True Positives |
| TN | True Negatives |
| ZGT | Ziekenhuisgroep Twente |

# TABLE OF CONTENT

# INTRODUCTION

Due to the ageing population, demand for intensive care is increasing (1). As a result, critically ill patients will be transferred to the general surgical ward earlier. These critically ill patients are often complex cases and have a higher chance of developing complications that can eventually result in mortality. Especially patients who undergo complex upper gastrointestinal surgery and geriatric patients with hip fracture were identified to be vulnerable for deterioration after surgery. Early detection of deterioration in these patients is critical in preventing major complications that can result in impairment, function loss, longer days spend in the hospital, higher hospital costs and even mortality (2-5).

In an attempt to lower patient mortality rates and better patient outcomes, hospitals have introduced the Rapid Response System (RRS). These systems are designed to help patients who require acute medical care and to prevent adverse patient outcomes (6).

It has been shown that several hours before an adverse event, physiological deterioration can be detected based on vital signs (7-8). Hospitals use these physiological changes and to activate the rapid response systems (RRS). Changes in vital signs have shown to be a powerful tool in predicting adverse events when measured regularly (9). Nurses are expected to measure and report vital signs at least once every nurse shift, or more in case deterioration of the patient is suspected. Nurses use an instrument called the modified early warning score (MEWS) to assess whether vital signs are deviating from their normal ranges. Unfortunately the current method of measuring and recording vital signs remains suboptimal, resulting in an underestimating physiological deterioration in patients (9-17). The time interval between measurements is often large and vital sign fluctuate, which means measurements only show a momentary event and can, therefore, be misleading (11).The study by McGain, F et all, shows that the completeness of measurements of vital signs in the three days after major surgery was only 17% (17). Another study showed that 81% of the patients who had died due to adverse advents could have been identified early on, by measuring and recording vital signs regularly (15).

However, not only the change in vital signs has shown to be valuable in predicting adverse events. Research has shown that nurse observations and their intuition have an even higher predictive power than deviating vital signs (18-19). Kim, Y et all (2015) even show that 46.8 % of the patients still had hardly had any physiological changes eight hours before a cardiac event (16), emphasizing the importance of nurse intuition. Despite the predictive power of nurse intuition, practice shows that the likelihood of RRS activation is higher when vital signs are deviating a lot, and severe or abrupt changes occur, compared to subtle clinical changes that are often picked up by the nurse (17). This implies that the nurse intuition alone is not enough to activate the system and should, therefore, be combined with information about vital signs. A study combining vital signs and nurse worry to predict adverse events in patients shows that this increases the predictive value of vital signs and Nurse Worry alone (19). This emphasizes that Nurse Worry is indeed essential in predicting adverse events and therefore, should be used more often. Nevertheless, this does not solve the issue of inconsistent measurements and incomplete vital signs data.

Continuous vital monitoring potentially resolves the shortcomings of the current practice and already shows promising results in the early detection of deterioration (2,21). With continuous vital monitoring, nurses will not have to measure vital signs manually, which reduces workload. However, it has not yet been established whether combining continuous vital sign monitoring with nurse worry will increase the detection of deteriorating patient condition in continuous data as well, or if only monitoring vital signs alone is enough to detect adverse events.

Therefore the goal of the study is to firstly, investigate the predictive value of the MEWS that nurses use to detect deterioration and investigate whether adding nurse worry indicators add value to the earlier prediction of deterioration. Secondly, it will be investigated whether continuous vital sign monitoring combined with nurse worry increases the detection of deterioration in surgical patients and also whether the deterioration can be detected earlier.

## RESEARCH QUESTIONS

To achieve the goals of this study, the following research questions and sub-questions will be answered.

1. *How do manually obtained nurse Worry and MEWS perform in predicting adverse events in patients at the surgical ward at the ZGT hospital over six months?'*
    - What are the diagnostic values of the manually obtained nurse worry and MEWS as composite scores in high-risk surgical patients?
    - What individual nurse worry indicators and manually obtained vital sign observations are good predictors when predicting an adverse event in high-risk patients of the surgical ward?
    - What is the average time between the first detection by manually obtained nurse worry /MEWS and the confirmation of an adverse event by current clinical diagnosis?

2. *Will combining the best predictors of the manual vital sign observations and nurse worry indicators increase the performance of predicting adverse events in continuous monitoring compared to the traditional MEWS scored on continuous data?*

- What is the predictive performance of the traditional MEWS in predicting adverse events when used on continuous vital signs?
- What is the predictive performance when the best predictors of manual vital sign observations and nurse worry indicators are combined into one model and scored on continuous data?
- How do different time intervals for the running average affect the performance of the model on the continuous data?
- What is the estimated time gain of the first detection of an adverse event by the combined model compared to the Nurse worry and MEWS on manually obtained data?

# THEORETICAL BACKGROUND

## 1. RAPID RESPONSE SYSTEM

Rapid response systems (RRS) are introduced by hospitals to increase safety for patients in hospitals (6). The RRS describes the whole process that is involved in the recognition and management of deterioration in patients (21). The introduction of the RRS was a result of different studies showing that half of all severe complications were preventable (22-23). One of the contributing factors was inadequate communication between professionals and inexperienced clinicians. The systems consist of detection of deterioration and triggers that set of response (afferent limb), a rapid response team (efferent limb), an administrative limb and quality improvement or governance limb (figure 1).

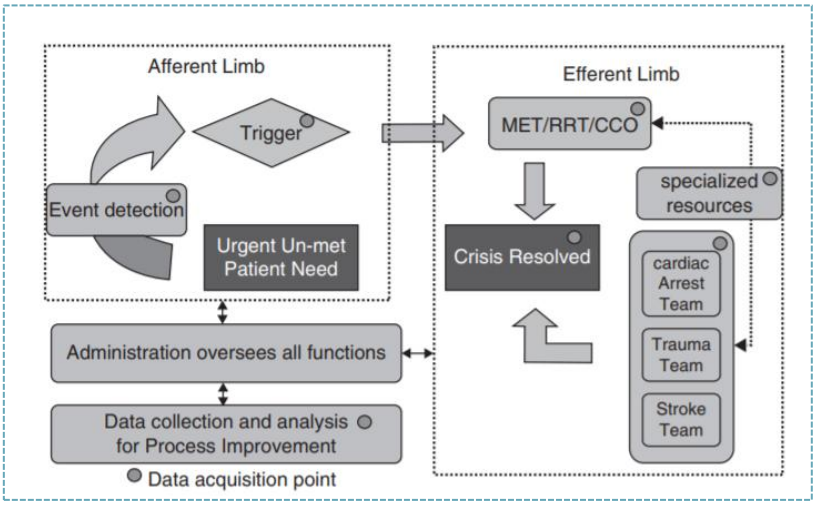



**Figure 1**. Overview of the rapid response system components. Figure copied from ludikhuize et all (24).

### *Afferent limb*

The afferent limb is the most important limb in the RRS (24). This component deals with detection of deterioration and sets off a response. Nurses working on the general wards are responsible for assessing patients, recognizing potential deterioration and to call for help when needed. The RRS protocol in the Netherlands differs a bit from other countries. In the Netherlands, the primary responsible physician is called in case of deterioration, instead of the rapid response team (24). The modified earlier warning score (MEWS) is used to help nurses detect deterioration. MEWS is a scoring system based on vital functions and patients observations (elucidated in H2). The outcome of the MEWS can be used to set of response when needed. However, not only objective measures can set off a response. Literature shows that the subjective 'worry' of nurses was the most used reasons for calling (19).

### *Efferent limb*

The efferent limb consists of personnel and equipment that is used to help the patients in critical situations (12). An essential part of the efferent limb is the rapid response team (RRT). RRT is a group that usually consists out of clinicians and ICU personal that provide rapid care when patients show signs of deterioration. The RRT is trained to help patients in critical situations.

### *Administrative and quality improvement limb*

The administrative limb is responsible for guiding personnel and equipment resources (12). The administrative limb cannot function properly without a quality improvement limb. This limb ensures quality is assessed, and new possibilities of improvement are identified. Continuous monitoring is one of the new options that are evaluated. Continuous monitoring can be used to make the RRS more effective.

The afferent limb is the most crucial component of the RRS but also most sensitive to errors (12). Without the detection of deterioration, there will be no response. The responsibly of detections lies with the nurses. These nurses make use of the MEWS score and their own experience (sense of worry) to call for help.  MEWS scores are often not conducted frequently enough, are incomplete or are not documented the right way (13-14). One of the main reasons for these shortcomings is the lack of time (high workload) nurses experience during their shift; as a result, signs of deterioration can be missed.  This implies that continuous monitoring could potentially be beneficial to nurses to detect deterioration and reduce workload.

## 2. MODIFIED EARLY WARNING SCORE (MEWS)

The earlier warning score (EWS) was first proposed in 1997 by Morgan, R et all(7). The EWS was developed as a response to accumulating evidence that suggested that patients in hospital wards showed physiological signs of deterioration several hours before a severe adverse event like cardiac arrest (7-8) The EWS is based on these physiological changes. Nurses use the EWS in hospital environment to detect signs of deterioration, where it provides validation to act on abnormalities (25). In 2000 the EWS score was changed to MEWS when new observations were added, such as oxygen saturation and urine production (ml/hour). These new observations were added to increase the predictability of adverse events (26).

### *MEWS scoring*

The MEWS uses patient vital functions to calculate a total score (table 1). The MEWS score is conducted at least every nurse shift (8 hours). Interventions based on the MEWS score can differ per hospital. In the Netherlands, it is recommended to call the primary responsible physician when a MEWS of ≥3 is recorded. With a MEWS score of ≥3, the physician is obligated to make a treatment plan within 30 min. When the MEWS score is ≥7 the physician should immediately see the patient.

| Score | 3 | 2 | 1 | 0 | 1 | 2 | 3 |
|---|---|---|---|---|---|---|---|
| **Respiratory frequency (breaths/min)** | | ≤8 | | 9-14 | 15-20 | 21-29 | ≥30 |
| **Heart rate (beats/min)** | | ≤40 | 41-50 | 51-100 | 101-110 | 111-129 | ≥130 |
| **Systolic blood pressure (mmHg)** | ≤70 | 71-80 | 81-100 | 101-200 | | ≥201 | |
| **Consciousness** | | | Confused | Alert | Responds to voice | Responds to pain | No response |
| **Body temperature (°C)** | | ≤34.9 | | 35-38.4 | | ≥38.5 | |
| **Urine production (ml/hour)** | | ≤20 | 21-39 | ≥40 | | | |
| **Oxygen saturation (%SpO2)** | ≤91 | 92-93 | 94-95 | ≥96 | | | |

**Table 1 Modified early warning score**

### *MEWS in clinical practice*

As was highlighted in the introduction, several hours before an adverse event, physiological deterioration can be detected (7-8). One study showed that a MEWS score ≥4 has a sensitivity of 75% and specificity of 83% in predicting the development of adverse events (9). This indicates that the MEWS score is a powerful tool in predicting adverse event when used regularly. However, a study looking into predicting cardiac arrest showed that 46.8% of the patients still had a low MEWS (≤2), 8 hours before the event (16).

The adherence to MEWS protocols is studied extensively. Incomplete MEWS score and poor compliance with the MEWS protocol is described in several studies (9-17). As a result, patient's physiological deterioration is generally underestimated. One of the studies shows that the completeness of the MEWS scores in the first three days after major surgery was only 17% (17). They described the respiratory rate to be missing the most (15.9%). Furthermore, one of the studies showed that 81% of the patients who had died due to adverse advents could have been identified early on by recording the MEWS frequently(15). The most common reasons for poor compliance are:

- Workload
- Fear of criticism
    - Nurses experience fear of not having skills or knowledge to provide quantifiable information to set-off response.
    - Low levels of confidence.
- Wrong calculation of score, results in both over and under-reporting of vital signs.
- Scarcity of equipment
- Communication between nurses.

So, MEWS has shown to be an asset for nurses to predict deterioration. Moreover the MEWS is used by nurses as clinical validation when calling for help. However, the use of the MEWS is suboptimal, resulting in missing signs and occurring of preventable events. The most important reasons for nurses to be non-adherent is having work pressure and fear of criticism for not being able to provide quantifiable information. This last reason implies that there is a need for nurses to have valid quantifiable information without an increase in workload.

## 3. NURSE WORRY INDICATORS

The nurses working on medical wards are a critical resource in detecting deterioration of patients. They see patients most often and can identify slight changes in the condition of their patients. Literature shows that nurses can recognize deterioration in patients before the current instruments like MEWS detect it (18-19). It turns out that intuition plays an essential role for nurses to identify deterioration and use vital signs to validate their feeling (27).

The importance of nurse worry has not gone unnoticed. In some hospitals "nurse worry," is scored as an additional point on the MEWS instrument. Several studies point out that nurse worry is also the most used criteria to activate the Rapid Response team (28). However, in practice, nurse worry is not always valued properly (15,19,29). Furthermore, nurses experience barriers to call when they feel worried about a patient (13). Fear of criticism or lack of objective data to justify their call, are the main barriers experienced by nurses when deciding whether or not to make the call (13-29). A systematic review has identified what underlying signs and symptoms were most often the basis for nurse worry (19). These sign and symptoms were used to construct the Dutch-Early-Nurse-Worry-Indicator-Score ( DENWIS.); an instrument that nurses could use to describe their worry when activating the RRT if needed. Despite promising results of DENWIS, these nurse worry indicators have not yet been fully implemented in Dutch hospitals.

| Indicator | Underlying signs and symptoms |
|---|---|
| Change in breathing | Noisy breathing and/or short of breath and/or unable to speak full sentences and/or use of accessory muscles |
| Change in circulation | Colour changes and/or clammy and/or coldness and/or impaired perfusion and/or oedema |
| Rigours | Rigours |
| Change in mental state | Lethargic and/or confused |
| Agitation | Restless and/or anxious |
| Pain | New pain and/or increasing pain |
| No progress | No progress and/or abdominal distension and/or nausea and/or bleeding and/or dizzy/fall |
| Patient indicates | Not feeling well and/or feeling of impending doom |
| Subjective nurse observation | Change in behaviour and/or doesn't look good and/or look in the eyes |

**Table 2.** The nurse worry indicators of the DENWIS instrument

# 4. CONTINUOUS MONITORING

The afferent limb of the rapid response systems have shown to be the most important limb of the whole system. Without detection of deterioration there will be no activation of the rapid response team. Unfortunately, in the current practice the afferent limb shows to be lacking. Study shows that a large part of the patients who experience an adverse event could have been detected if monitored more frequently (15).

For these reasons continuous monitoring is gaining more popularity and has already showed some promising results. Duus CL, Aasvang EK, Olsen RM, et al (2018), show that severe micro events based on parameters Spo2, HR en RR were detected significantly better by continuous monitoring than using the normal EWS. Especially hypoxemia was detected more consistently through continuous monitoring. Hypoxemia can inflict organ damage, such as to the heart and brain when not treated timely (2). Continuous monitoring is already shown to be feasible in the critical care setting, where it may provide better patient outcomes and be more cost effective than MEWS (32). Another study showed that the majority of patients, relatives and nurses were positive about continuous monitoring (20).

## ***Sensors***

There are many different sensors on the market that advertise wireless continuous monitoring in a hospital setting. However there are still very few studies that show how well these sensors perform. A system that has been tested was the early sense monitoring system. This systems measured heart rate, beat-to-beat fluctuation and bed motion (46). The bed motion sensor of the early sense monitoring system is placed under the mattress, which means it only works when the patient is in their bed. However, results of the study show that by using this measurement system the length of stay and intensive care days for transferred patients significantly decreased (47). There are several sensors that are place directly on the patient's skin and can be worn at any time. This means the patients can walk around and still be monitored (46). The biggest challenge in continuous monitoring is to ensure patient data is safely connected and stored (46). Furthermore accuracy, reliability and battery life are design specifications

that need to be tested to determine if these sensors are good enough to be used on the generals wards.

An example is the Isansys Lifetouch sensors. These sensors are also used in this study. The Isanys Lifetouch is a wireless biosensor pad that is placed on the thorax (30). The sensor can continuously analyse ECG signals and can generate data from which heart rate, respiration rate and Heart rate variability (HVR) can be calculated. The Insanys Lifetouch also contains an accelerometer that can measure information on patient orientation, activity and motion. The Isansys LifeTemp is placed on the axillary skin (31). Temperature measures are taken every 10 seconds and updated each minute. The Nonin 3150 wristOx2 is used to monitor oxygen saturation. All the devises are connected to the Patient status engine, where all the data is continuously collected. The collected information is encrypted to and continuously send to the patient digitalisation by Bluetooth. The patient status engine transfers the data to a local server ia Wi-FI and is stored in a secured database which is located in the hospital (figure 2)

Several different technologies can be used to monitor patients continuously (Table 3).

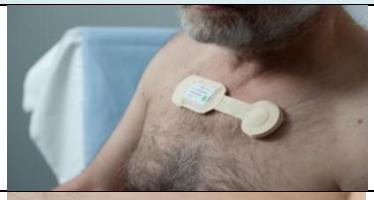
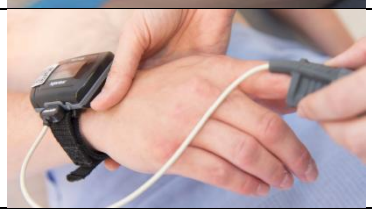
| Measurements | Sensors | |
|---|---|---|
| Heart rate, respiratory frequency and accelerometery | Isanys Lifetouch |  |
| Skin Temperature | Isansys LifeTemp |  |
| Oxygen saturation | Pulse oximeter Nonin 3150 WristOx2 |  |

**Table 3.** sensors used to monitor patients

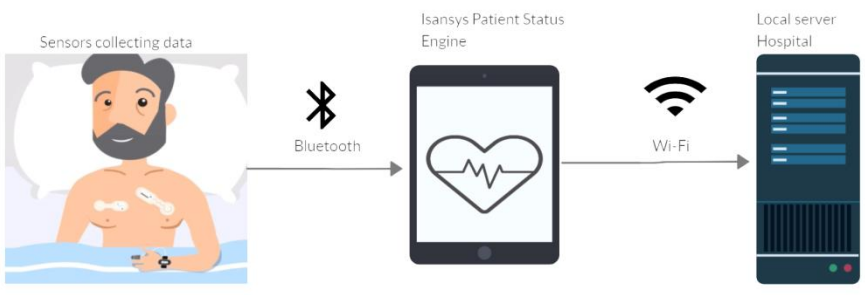



**Figure 2.** Wireless data transportation

## 5. CLAVIEN DINDO CLASSIFICATION

For a long time, there was no clear consensus about how to classify the severity or grade of surgical complications (46). In 1992 the first approach was proposed to rank the complication based on the therapy that was used to treat it and the type of negative outcome after surgery (46). The three different types of adverse outcomes after surgery that are used to classify are:

- Complication
- Failure to cure
- Sequela

In 2004 the ranking was reviewed, and a new 5-scale classification system was developed(table 4). This scale was only based on the type of therapy that was needed to treat the complications. "The rationale to preserve this approach was to eliminate subjective interpretation of serious adverse events and any tendency to down-grade complications, because it is based on data that are usually well documented and easily verified"(46).

Complication grade after surgery is classified by the following classification (table 4).

| Grades | Definition |
|---|---|
| Grade I | Any deviation from the normal postoperative course without the need for pharmacological treatment or surgical, endoscopic and radiological inventions |
| Grade II | Acceptable therapeutic regimens are drugs as antiemetics, antipyretics, analgesics, diuretics, diuretics and electrolytes and physiotherapy. This grade also includes wound infection opened at bedside |
| Grade III | Requiring surgical, endoscopic or radiological intervention |
| Grade III-a | Intervention not under general anaesthesia |
| Grade III-b | Intervention under general anaesthesia |
| Grade IV | Life-threatening complication (including CNS complications)* requiring IC/ICU management |
| Grade IV-a | Single organ dysfunction (including dialysis) |
| Grade IV-b | Multi-organ dysfunction |
| Grade V | Dead of a patient |

**Table 4**. Clavien Dindo classification

# RESEARCH DESIGN AND METHODOLOGY

This study is part of a more prominent study that is conducted by Ziekenhuisgroep Twente (ZGT) and the University of the Twente. The goal of this study is to investigate to what extent continuous monitoring vital sign monitoring can improve the number of events detected and the time to detect adverse in high-risk surgical wards patients.

## RESEARCH DESIGN

An exploratory study is conducted with secondary data that was provided by the Movisign research group. In the first part of this thesis, the predictive value of the MEWS and Nurse worry as composite scores is investigated. This provides insight into how well both models can predict adverse events separately in current practice. These outcomes are also used to compare the predictive value of the new model. Then, the separate variables that make up the nurse worry, and the MEWS are investigated to understand how they contribute to the model. Based on this information, a new model is constructed that combines the most predictive vital signs and nurse worry indicators to predict adverse events in patients. The new model is tested on continuous vital signs observations for different thresholds. The threshold with the most predictive power is finally compared to the current MEWS and nurse Worry. The traditional MEWS will also be tested on the continuous data to create baseline measurements (figure 6).

## STUDY POPULATION

Participants were patient aged >18 undergoing esophageal and gastric resection admitted to the surgical ward for postoperative care and patient >70 undergoing hip fracture surgery and admitted to the surgical ward for pre- or postoperative care. These patients groups were chosen due to often high risk for developing adverse events. The patients were recruited at ZGT hospital in Almelo for over six months. Inclusion and exclusion criteria were:

*Inclusion criteria*

- >18 years undergoing elective surgery for resection of malignant tumors of the upper gastrointestinal tract and admitted to the surgical gastrointestinal ward for postoperative care

- >70 years undergoing surgery for a hip fracture and admitted to the geriatric traumatology

*Exclusion criteria*

- Suspected or diagnosed delirium, cognitive impairment, or dementia
- Known allergy/skin irritation to the adhesives used in the sensor patches.
- Broken or irritated skin in the sensor placement area
- Disease caused by prions (i.e. transmissible spongiform encephalopathy)
- Contact isolation
- Implanted medical devices, such as a cardioverter-defibrillator or pacemaker
- Incapacity to decide on study participation

## DATA COLLECTION

All the data used was previously collected for the Movisign study. The data was stored in a database. Data about the patients and health professionals that could lead to the identification of the subjects were coded, so all data was anonymous and untraceable to a person. The data that was part of the standard care was extracted from the electronic medical records. A checklist was used to collect information about nurse worry. Nurses were asked to fill in a worry checklist at least once every nurse shift or more in case of Worry. In the list, nurses were asked if they were worried and what the reason for their concern was. Once they completed the checklist, they were unable to access the filled-in list. The whole checklist can be found in appendix (A). The exact number of different nurses is unknown. Only information about years of work experienced was collected. Continuous monitoring data was collected by sensors that were provided by Isansys Lifetouch. The continuous measurements were not visible for nurses, physicians, or patients. All the data collected by the sensors is sent to the Patient Digitations engine via Bluetooth. The data was then transferred to a local server via Wi-Fi and stored in a database within the hospital.

Different types of data was collected and used in this study. In total, four different datasets with information were used. All the variables can be found in the Appendix (B). Most important variables are explained in the following paragraphs.

### Patient characteristics

all the measurements, accept for adverse event data, is recorded at the time of inclusions by a healthcare professional. Adverse events are that are recorded were confirmed by standard protocol throughout the study.

> Adverse event : ''an undesirable experience occurring during the study''

An adverse event is included when it meets a set of criteria. First, it is a postoperative complication classified as Clavien Dindo class of II (table 4), diagnosed according to standard protocol. Furthermore, the complication should have started at least 24 hours after the start of the measurement period. Besides adverse events, the timing of the event and the type of event is collected

> *Adverse event timing*: '' the time that adverse event was detected and confirmed according to standard protocol''.

> *Type of event:* '' the type of complication that is detected and confirmed according to standard protocol''

### Nurse worry

Nurse worry and the reasons to worry (NW indicators) were taken from the nurse worry checklist and used in the analyses. All variables were dichotomous( Yes/No) and are based on DENWIS(19).

- *Worry*: worry about the patient?
- *Change in breathing:* Patients breathing is changing—for example, noisy breathing, short of breath, unable to speak full sentences and use accessory muscles.

- *Change in circulation:* Skin colour changes, clammy feeling, coldness, impaired perfusion or oedema.
- *Temperature*: the patient has rigors.
- *Change in mental state:* the patient is lethargic or confused.
- *Agitation:* Patient is restless or anxious
- *Pain:* Patient is experiencing new pain or increasing pain.
- *Unexpected course:* the patient is showing no progress, nausea, bleeding, fainting, or falls.
- *Patient feeling:* Patients indicates not feeling well or feeling of pending doom.
- *Family feeling:* family indicates something is wrong with the patient. Change in behaviour, change in attitude, not looking well or change in look in the eyes.
- *Nurse feeling:* Nurse indicates a change in behaviour; the patient doesn't look right or change in the look in the eyes.

### *Manual vital sign observation*

Manual vital sign observations are conducted as part of the standard care at the surgical ward. Vital sign observations were conducted at least once every nurse shift. In the case of deviating signs or 'nurse worry', nurses can choose to increase the frequency of manual vital sign observations. The MEWS is calculated based on the vital signs.

- *Heart rate*: calculated by beats per minute. Numerical data.
- *Respiration rate*: Calculated by breaths per minute. Numerical data.
- *Body temperature:* unit used is °C. Numerical data.
- *Oxygen saturation (Spo2):* Oxygen saturation measures oxygen levels in the blood. It is expressed in percentages were 100% is the maximum. Numerical data.
- *Systolic blood pressure:* unit is used is mmHg. Numerical data.

### *Continuous data*

Continuous data is measured on a one-minute interval. The type of data collected is the same as in the manual sign observation. However, systolic blood pressure is not measured in the continuous data

Both the manual vital signs observations and the continuous data had missing data (table 5).

| Missing data | Manual vital sign observation | Continuous data |
|---|---|---|
| Heart rate (beats/min) | 4.9% | 30% |
| Respiratory rate (breaths/min) | 45% | 30% |
| Body temperature (°C) | 7.8% | 24 % |
| SpoO2 (%) | 8.7 % | 58% |
| Systolic blood pressure (mmHg) | 6.3 % | - |

**Table 5.** Percentages of missing data

Manual vital sign observations file contained a total of 1135 observations. The continuous data file contained a total of 535118 observations. In the manual vital sign observation file, all data is imputed with last observation carried forward (LOCF). Missing data in the continuous data file data imputed with last observation carried forward as well. Data is imputed to increase the amount of available data. However in the continuous data file, only the measurements which are missing for less than an hour continuously will be imputed. It is to be excepted that imputing data that is missing for longer than an hour will results in significant bias.

Most data was missing due to connection or recording failures. Literature was searched to compare the amount of missing data in other study. One study shows significantly less missing data (20). However another study showed similar amount of missing data (48).

For part 2 of the thesis, the Nurse Worry file and the manual vital sign observation file were merged. This was done matching the Nurse Worry with the manual vital sign observation that was taken in the same shift. When Nurse Worry was expressed, but there was no manual vital sign observation within the same shift, then it was matched with the first manual vital sign observation after the Nurse worry was expressed.

## *DATA ANALYSIS*

### *Baseline characteristics*

Baseline characteristics of the patients with an event (event group) and without event (control group) will be compared. Nominal variables are expressed in frequencies (%) and are analyzed with the chi-square test. The mean is calculated for the continuous variables. The student T-test or Mann-Whitney U test will be used to calculate the standard deviation depending on the distribution of the continuous variables.

## PART 1. DIAGNOSTIC VALUES NW AND MEWS

Frequencies of expression of NW, NW indicators and averages of manually obtained vital signs observations of the event group and the control group will be compared, similar to the baseline characteristics.

The diagnostic values of the NW and the manually obtained MEWS as composite measures will be investigated by calculating sensitivity, specificity, positive- and negative predictive values, and the Area Under the Curve (AUC) value.

The detection of an adverse event by the MEWS is defined as; a MEWS of ≥3 at least once after start of the study and before an adverse event occurred, according to the complication timing in the dataset. The complication timing is the detection and confirmation of an adverse event according to standard protocol. Detection of an event by NW is defined as; Worry expressed least once after the start of the study and before an adverse event occurred according to the complication timing. The calculations will be done by creating a 2x2 table with four different categories: True positives (TP), false positives (FP), true negatives (TN) and false negatives (FN).

Sensitivity is the percentage of patients who were classified as positives and who indeed experienced an adverse event. Specificity, on the other hand, shows the percentage of patients who were classified as negatives and indeed experienced no adverse event.

Furthermore, positive and negative predictive values will be calculated. These values represent the diagnostic power of a test (34). The positive and negative predictive value of NW depends on the prevalence of patients who developed an adverse event. In the calculation, it is assumed that the event rate in the current sample reflects the prevalence of developing adverse events for these patients groups.

Lastly, the Area Under the Curve (AUC) will be calculated using SPSS. The AUC shows the distinguishing abilities of a test with a certain threshold. This value can be used as an overall measure that tells us how well the test classifies the patients in the right category. To evaluate the distinguishing abilities of the tests, the rule of thumb will be used (35):

- 0,5 → no discriminative power
- 0,5-0,7 → poor discriminative power
- 0,7-0,8 → acceptable discriminative power
- >0,9 → outstanding discriminative power

## PART 1B. PREDICTIVE VALUES

### *Logistic regression*

The separate manually obtained vital signs that make up the MEWS and the independent indicators of NW will be analyzed in a univariate logistic regression analysis to investigate their predictive power. Data before the confirmation of an adverse event is used in the event group. In the control group, all data was used.

Next, a multivariate analysis of the manually obtained vital signs will be conducted to investigate which combination of variables has the highest predictive power. The same will be done for the indicators that make up the NW. In both multivariate analyses, the backward selection will be used to select the variables that have the most predictive power. The most predictive variables of the manual vital signs observations and the nurse worry indicators will be combined in the next step.

For the logistic regression analyses, the vital signs are categorized from 0 to 3 based on what they would score on the MEWS. The categorized vital signs will be treated as continuous in the logistic regression. The nurse worry indicators are dichotomous of nature and are assigned either 0 or 1.

Due to small nurse worry data (60 observations), higher p-values threshold are chosen to state that a variable is important. For both the univariate and multivariate analyses variables are considered to be of importance with a p-value of 0,30. Variables that have less than six observations will not be used in the analysis. Furthermore, variables will be tested on correlation if variables that are highly correlated (arbitrary threshold of 0.6), one of the variables will be taken out of the analysis (36).

| Score | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| **Respiratory frequency (breaths/min)** | 9-14 | 15-20 | 21-29 | ≥30 |
| | | | ≤8 | |
| **Heart rate (beats/min)** | 51-100 | 101-110 | 111-129 | ≥130 |
| | | 41-50 | ≤40 | |
| **Systolic blood pressure (mmHg)** | 101-200 | 81-100 | ≥201 | ≤70 |
| | | | 71-80 | |
| **Body temperature (°C)** | 35-38.4 | | ≥38.5 | |
| | | | ≤34.9 | |
| **Oxygen saturation (%SpO2)** | ≥96 | 94-95 | 92-93 | ≤91 |

**Table 5.** Ranges of vital signs in mews scores

## PART 2AB. NEW WARNING SCORING INSTRUMENT FOR CONTINUOUS MONITORING

The most predictive values of the manually obtained vital signs observations and nurse worry indicators of the previous step will be placed together in multiple logistic regression. Backward selection with a p-value threshold of ≥0,30 will be used to come to best predicting model. The beta coefficients of the end model will be used to calculate scores for the variables in the new continuous warning score model (CWS). Hemalkumar et all, (2016), showed in their article that developing a scoring system based on the beta coefficients is an accurate way to develop a risk scoring system. Beta coefficients will be multiplied by ten and round to integer. This will be done to discriminate between the predictive values. This method of weighing the beta coefficients showed the most significant net reclassification improvement compared to the baseline model. The new continuous warning scoring model (CWS) will work similar as the traditional MEWS.

When making a predictive model, the size of the dataset influences the outcomes of the model. In a smaller dataset, the biggest problem is over fitting. Due to the often few events or observations, the variance between the measures is significant. When this happens, the model predicts well on the training data but often shows poor outcomes when external validation on a new data set is conducted (pavlou 2015). To reduce the over fitting of the model one should look at the number of events per variable (EPV) (38). The EPV is a ratio of the number of events divided by the number of variables in the risk model. It has been suggested that an EPV of 10 or more could reduce the problem of over fitting. The problem with over fitting is that the model will underestimate low-risk patients and overestimate high-risk patients, which leads to a waste of resources. Unfortunately, if this rule of thumb of an EPV ≥10 would be used in this dataset (event N=13), then only 1 to 2 variables can be used to build a predictive model.

To overcome the problem of over fitting and to reduce the variables considerably, Pavlou et al. 2015 suggest using a penalized logistic regression model. The idea behind the penalized regression is that it will shrink the regression coefficients of the model, making them less sensitive to the

considerable variation due to the low number of events. The shrinking will be done by placing a constraint on the regression coefficients. Different penalization methods can be used. The ridge method is the preferred method as this will make sure that all preselected variables can stay in the model (39). The restrain in the ridge method is derived with the following formula:

$L = \lambda \sum \beta^2$ … (41)

The value given for $\lambda$ is determined by the researcher and determines the size of the restrain. The most optimal value for $\lambda$ can be investigated by cross-validation. The most common chosen approach is de 10-fold cross-validation. The complete analysis will be conducted using R.

To investigate whether the ridge regression method will indeed improve the new model, two scoring models are made. One model uses the scores based on the coefficient of 'normal' logistic regression, while the other model is based on the beta coefficients of the ridge regression, the CWS2. Outcomes of both models are compared.

## PART 2C. TEST THE MODEL ON CONTINUOUS DATA

The new scoring models will be tested on the continuously obtained vital signs data. The detection of an event is defined in the same way as the original MEWS. In the normal MEWS, the threshold a MEWS of ≥3 is considered to be the optimal score to detect deterioration and trigger a rapid response. To determine a good rapid response trigger threshold for the new scoring model, sensitivity/specificity, PV+, PV- and the AUC will be calculated for every possible score. Specificity and negative predictive value will be of most importance since false negatives lead to missed deterioration and possible severe health consequences for the patient. However, sensitivity should also be at least as good as the old MEWS / NW. A lower sensitivity will lead to more false positives and can potentially lead to wasted time and recourses. Finding the best threshold score will be done for both CWS and CWS 2 risk scoring instruments.

Furthermore, the events that are not detected in the continuous data will be analyzed to get insight into what kind of adverse events are missed. The normal MEWS will also be tested on the continuous data for baseline comparisons. The MEWS will be tested and analyzed the same as the CWS model; only the scores of the MEWS will be used. The MEWS on continuous data is called the CMEWS.

### *Continuous data*

As was described before, the continuous data only contains heart rate, respiratory rate, saturation (SPo2), and auxiliary temperature recorded at one-minute intervals. To test the whole model, manually recorded nurse worry and systolic blood pressure are used. In the analysis, the moment systolic blood pressure deviates, or nurse worry is expressed, the scores assigned to these variables are carried forward for four hours. If no new nurse worry or systolic blood pressure is noted during these four hours, their respective values are assumed to be within an acceptable range and will be score 0.

Furthermore, the continuous data of the sensors is incomplete. Scoring for missing values are handled in the following way: When two or more of the four sensor measures have been missing continuously for the previous hour, no score is recorded at the considered time step. When the measures are not present for less than an hour, then the last observation carried forward method is used. When only one variable has missing data for over an hour, the recordings of the remaining variables will be used to score the patient at that time step (figure 3).

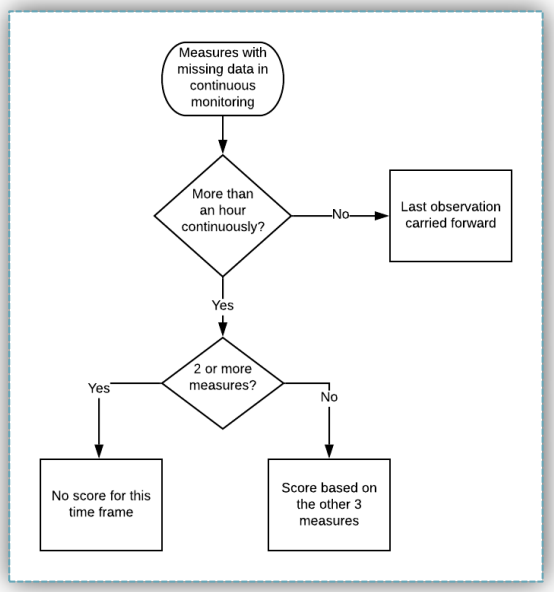



**Figure 3.** Scoring for missing data continuous monitoring.

As was described, the measurements are taken at one-minute intervals. However, if we would apply the scoring system to the one-minute interval data it will probably give a lot of false positives due to short term fluctuations due to, for example, increased heart rate due to walking. To detect deterioration and reduce 'noise', the measurements will be smooth out by taking the moving average of the vital sign different time intervals. The CWS models are tested on 5-10-30-60-120-240-480 minute intervals to see which time range best suits each model. For all different time intervals, the best threshold is calculated. The best threshold is chosen based on the highest combination of sensitivity/specificity.

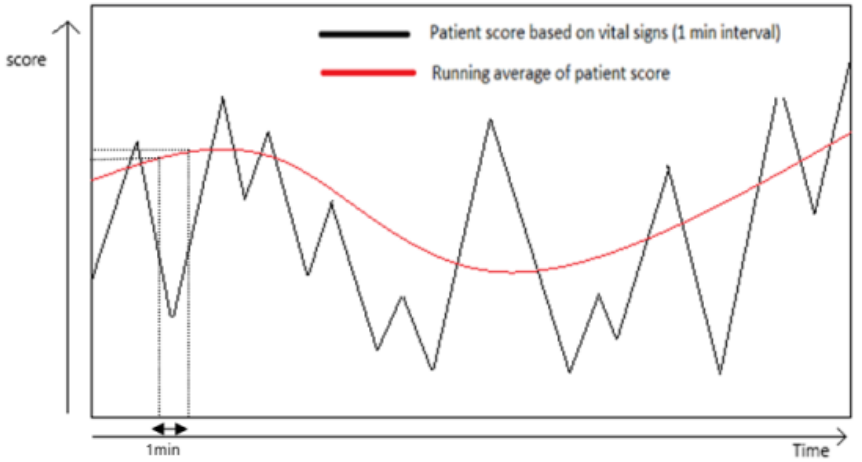



**Figure 4.** Running average score in continuous data.

## Detection time

The best threshold scores  of the CWS models will be used to calculate the detection time for 5-60-480 minute intervals. The detection time is defined as the time between the first time the threshold score occurred in the continuous data and adverse event timing according to the current protocol (figure 5). These time intervals are chosen to see if there is a difference in detection time at a different time intervals.  Besides detection time, the type of events that were not detected at the optimal threshold for these 5-60-480 minute interval will be analyzed.  For the traditional MEWS and nurse worry on manual data and the MEWS on continuous data (CMEWS)  the detection time will be calculated similarly.  The threshold for  MEWS and CMEWS is ≥3.  The first time nurse worry was expressed is used to calculate detection time for nurse worry.

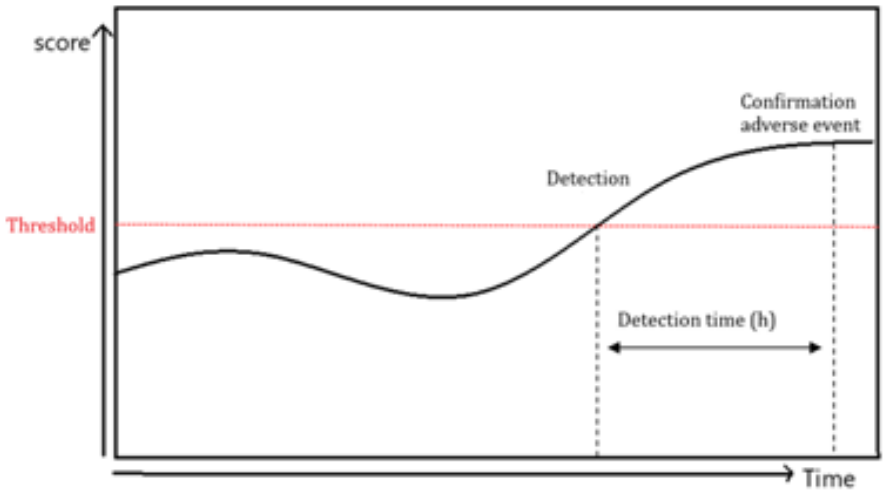
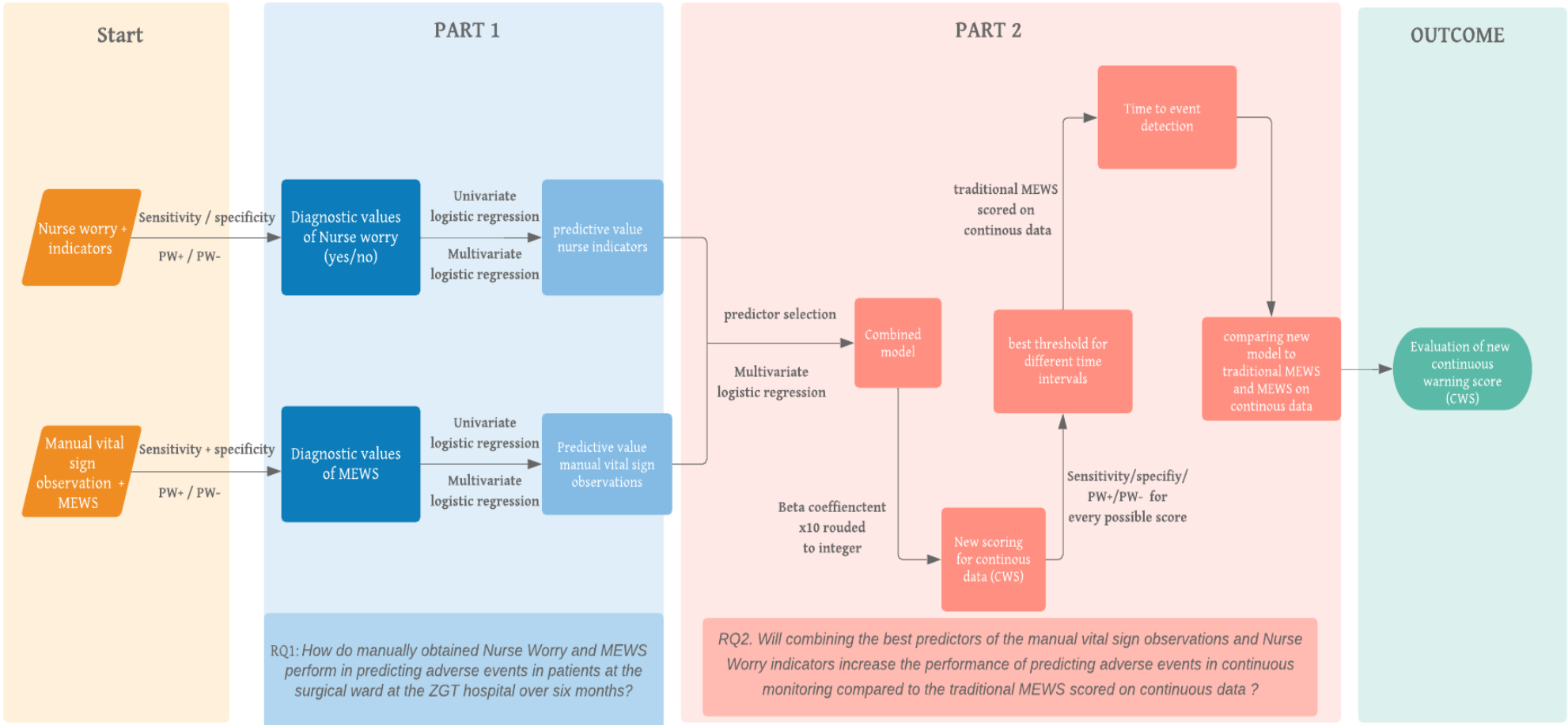



**Figure 5.** Detection time

**Figure 6** overview thesis



# RESULTS

## DEMOGRAPHICS

There were 60 patients in the file of subjects.

Of these 60 patients, 13 experienced an adverse event that had a Clavien Dindo class 2 and was confirmed during the experiment. The group with elective esophageal surgery had the most complications (n=6), followed by hip fracture (n=5) and elective gastric surgery (n=2). There were no significant differences in demographics between patients with and without an adverse event.

| Elective esophageal surgery | Elective gastric surgery | Hip fracture |
|---|---|---|
| Atrial fibrillation 2x<br>Pneumonia 4x | Pneumonia<br>Closed gastric surgery | Atrial fibrillation<br>Wound leakage<br>Urinary tract infection<br>Congestive heart failure<br>Myocardial infarct |

**Table 6** type of events registered in the patient groups.

| | Event group (n=13) | Control group(n=47) | p-value* |
|---|---|---|---|
| Female | 46.1 % (6) | 44.7% (26) | 0.755 |
| -Elective esophageal surgery | 46.1% (6) | 31.9% (15) | |
| - elective gastric surgery | 15.3% (2) | 21.3% (10) | 0.710 |
| - Traumatic hip fracture surgery & 70 years or older | 3.8% (5) | 46.8% (22) | |
| Average age | 73 (11.6) | 71.9 (13.2) | 0.790 |
| Average weight | 77.9 (22.1) | 77.4 (17.4) | 0.952 |
| Average height | 1.74 (0.11) | 1.74 ( 0.12) | 0.979 |
| Average ASA | 3.2 (0.77) | 2.5 (0.59) | 0.681 |
| Co morbidities | | | |
| - Cardiac | 38.5 % (5) | 29.8% (14) | 0.737 |
| - Vascular | 53.8% (7) | 48.9% (23) | 1.000 |
| - Diabetes | 0 | 21.3% (10) | 0.099 |
| - Pulmonary | 0 | 29.8% (14) | 0.027 |
| - Neuropsychiatric | 7% (1) | 31.9% (15) | 0.153 |
| - Gastrointestinal | 46.2% (6) | 29.8% (14) | 0.326 |
| - Urogenital | 46.2% (6) | 29.8% (14) | 0.326 |
| - Thrombotic | 0 | 2.1% (1) | 1.000 |
| - Neuromuscular | 15 % (2) | 6.4 % (3) | 0.427 |
| - Endocrine | 7% (1) | 21.3 % (10) | 0.295 |
| - Infective | 7% (1) | 6.4% (3) | 1.000 |
| - Other | 84.6% (11) | 68.1% (32) | 0.314 |

**Table 7.** Demographics of patient characteristics
*Difference significant at P<0.05

## PART 1A. DIAGNOSTIC VALUES

To get insight into how the MEWS and NW are performing separately in current practice, diagnostic values like sensitivity, specificity, PV+ and PV- of both instruments were calculated. Furthermore, differences between the event and control group were compared by calculating frequencies and averages of the components that make up the MEWS and NW.

### *Manual vital sign observations / MEWS*

In total, there were 1130 manual vital sign observations reported. Patients in the event group had, on average, 24 observations taken during their hospital stay. The patient without event had an average of 17 observations. The average MEWS of the patients during their entire hospital stay in the event group and the control group was 1.86 and 1.60, respectively. The mean for the separate manual vital signs observations are calculated (table 9). Heart rate is significantly higher in the event group (P=<0.001) and systolic blood pressure is significantly lower (p= 0.017).

| | Event group N=310 | Control N=820 | P-value* |
|---|---|---|---|
| Heart rate (beats/min) | 93 (0.9) | 86 (0.6) | <0.001 |
| Respiratory rate (breaths/min) | 17 (0.3) | 16 (0.2) | 0.015 |
| Temperature (°C) | 37,1 (0.03) | 37,1 (0.02) | 0.326 |
| SPo2 (%) | 95(0.1) | 95(0.8) | 0.150 |
| Systolic blood pressure (mmHg) | 124 (0.9) | 137 (1.8) | <0.001 |
| Average MEWS | 1,86 (1.6) | 1,60 (1.8) | 0.017 |

**Table 8**. Differences event vs control in manual vital sign observations.

*Significant at P ≤ 0,05

There was no significant difference between the event and control group in temperature and SPo2.

Diagnostic values were calculated for a MEWS of ≥3. All patients with an event had a MEWS of ≥3 preceding the event, resulting in a sensitivity of 100%. However, there were a lot of false positives (N=23), resulting in a specificity of 49 %. PV+ and PV- were 36 % and 100% respectively. The AUC was 0.6.

Furthermore, the first time a MEWS ≥3 was seen in the event group was on average 63.5 hours ( 2.6 days) before confirmation of the event.

### *Nurse Worry*

The nurse worry file contained information of 60 different patients. A total of 453 Worry checklist were conducted. The event group and the control group had 115 and 338 filled in worry checklists, respectively.

Worry was expressed 66 times. In the event group, it Worry was expressed 22% (n=25) of the time. In the control group, this was only 12%(n=41). This differences showed to be significant with a p-value of 0.021. An adverse event was confirmed in eight patients after nurse worry was expressed. There were only three patients who experienced an adverse event without nurse worry every being expressed. Some of the indicators for Worry showed significant differences between (table 10)

| | Event group (N=115) %(n) | Control group (N=338) | P-value* |
|---|---|---|---|
| Worried | 22% (25) | 12% (41) | 0.021 |
| Reason breathing | 9% (11) | 4% (14) | 0.028 |
| Reason circulation | 3.5 % (4) | 0.6% (2) | 0.019 |
| Reason temperature | 4.3% (5) | 2.4% (8) | 0.272 |
| Reason mental | 2.6%(3) | 1.8 % (6) | 0.580 |
| Reason agitation | 0% | 0.6% (2) | 0.408 |
| Reason pain | 5.2% (6) | 3.8 (13) | 0.526 |
| Reason unexpected course | 0% | 0% | - |
| Reason patient feeling | 7.8% (9) | 3.0 % | 0.024 |
| Reasons family feeling | 2.6% (3) | 0.3 % (1) | 0.022 |
| Nurse feeling | 8.7% (10) | 3.6% (12) | 0.027 |

**Table 10.** Differences Event vs control group nurse worry indicators in manually obtained data. * significant at p-value ≤0.05

Diagnostic values for nurse worry were calculated. Nurse worry alone performed worse than the MEWS. Sensitivity and specificity were 64% and 54%, respectively. Positively predictive value was 28% and the negative predictive value is 84%. The AUC was 0.54

First time Nurse worry was expressed in the event group was on averaged 65.8 hours (2.7 days) before confirmation of an event.

## PART 1B. PREDICTIVE VALUES

### Manual vital sign observations

First, a univariate analysis was conducted to find the variables that were significantly associated with the chance of experiencing an adverse event (p= ≤0.30). The univariate analysis was conducted for the manual vital sign observations which were categorized based on the MEWS (table 6).

Deviating heart rate, respiratory rate, Spo2, and systolic pressure significantly increases the chance of experiencing an adverse event (table 11).

When the heart rate is between the ranges of 101-110 or 41-50 BPM than heart rate received one point on the MEWS instrument (table 6). The analysis shows than the chance of experiencing an adverse event increase 1,6 times (p=<0.001)(table 11)

A heart rate value which scores two on the MEWS, which is equal to a BMP between either 111-129 or ≤8, shows to increase the chance of experiencing an adverse event with 2.6 times. Heart rate classified as MEWS 3 (>130 BPM) is showing 4.3 times increase in chance.

| variables | β | SE | P-value** | Odds ratio |
|---|---|---|---|---|
| Heart rate (1)* | 0.484 | 0.095 | <0.001 | 1.6 |
| (2) | | | | 2.6 |
| (3) | | | | 4.3 |
| Respiratory rate (1) | 0.249 | 0.087 | 0.004 | 1.3 |
| (2) | | | | 1.6 |
| (3) | | | | 2.1 |
| Temperature (1) | 1.111 | 0.178 | 0.528 | 0.3 |
| (2) | | | | 0.1 |
| SpO2 (1) | 0.148 | 0.082 | 0.071 | 0.9 |
| (2) | | | | 0.7 |
| (3) | | | | 0.6 |
| Syst. blood pressure | 0.604 | 0.252 | 0.016 | |
| (1) | | | | 1.8 |
| (2) | | | | 3.3 |
| (3) | | | | 6.1 |

**Table 11.** Univariate analysis manually obtained vital signs.
* MEWS score between brackets
**significant at p-value <0.30

A systolic pressure which is classified as a MEWS of 3 (<70 mmHG) is showing the highest increase of chance (OR=6.1)(p=0.016). Spo2 shows to decrease the chance of getting adverse (P=0.071). This result is noteworthy as one would expect these results to be positive. The MEWS was developed based on physiological changes in vital signs that were associated with experiencing an adverse event (23-25). Temperature is showing a decrease of chance as well; however, this result is not significant.

Next up, a multivariate logistic analysis was conducted to investigate the relationship between all the variables and the developed of an adverse event. With a backwards selection, the variables that best predict were selected (P-value ≤0.30). The following results were obtained.

All variable had p-values below ≤0.30 resulting in that all the variable are left in the model. There was no significant correlation between variables( threshold of 0.6).

Systolic blood pressure (P=0.007) and heart rate (p=< 0.001) seem to be the most important variables to increase the chance of getting adverse event when deviating based on the MEWS (table 12)

SPo2 (P=0.025) and temperature (P=0.001) are shown to be decreasing the chance with a MEWS of 1 with 0.8 and 0.5 times, respectively.

| variables | β | SE | P-value* | Odds ratio |
|---|---|---|---|---|
| Heart rate (1) | 0.661 | 0.119 | <0.001 | 1.9 |
| (2) | | | | 3.7 |
| (3) | | | | 7.2 |
| Respiratory rate (1) | 0.185 | 0.099 | 0.062 | 1.2 |
| (2) | | | | 1.4 |
| (3) | | | | 1.7 |
| Temperature (1) | -0.735 | 0.221 | 0.001 | 0.5 |
| (2) | | | | 0.2 |
| SpO2 (1) | -0.193 | 0.086 | 0.025 | 0.8 |
| (2) | | | | 0.7 |
| (3) | | | | 0.6 |
| Syst. blood pressure (1) | 0.686 | 0.256 | 0.007 | 2.0 |
| (2) | | | | 3.9 |

**Table 9.** Multivariate logistic regression manual vital signs

* significant at p≤0.30

## Nurse worry indicators

Similar to the manual vital sign observations, univariate and multivariate logistic regressions were conducted for the nurse worry indicators. The nurse indicators about circulation, agitation, family feeling and unexpected course were taken out of the analysis due to the small number of mentions (≤6 times).

| Indicators | β | SE | p-value* | Odds |
|---|---|---|---|---|
| Breathing | 0.895 | 0.418 | 0.032 | 2.5 |
| Temperature | 0.629 | 0.279 | 0.279 | 1.9 |
| Mental | 0.393 | 0.716 | 0.582 | 1.5 |
| Pain | 0.319 | 0.506 | 0.528 | 1.4 |
| Patient feeling | 1.024 | 0.473 | 0.030 | 2.8 |
| Nurse feeling | 0.951 | 0.443 | 0.032 | 2.6 |

**Table 13** univariate logistic regression NW.
* significant at P≤0.30

Of the remaining six variables, all were positive, suggesting that when worry about one of the variables was indicated the chance of developing an adverse event increases (table 13). However, breathing (P=0.032), temperature (p=0.279), patient feeling (p= 0.030) and nurse feeling(p=0.032)showed to be significant.

After backward selection (p<0.30), the multivariate regression analysis showed the following results (table 14). There were no significant correlations between variables.

| indicators | β | SE | p-value* | Odds |
|---|---|---|---|---|
| Patient feeling | 0.643 | 0.509 | 0.206 | 1.9 |
| Nurse feeling | 0.703 | 0.542 | 0.195 | 2.0 |
| Constant | -1.150 | 0.113 | 0.000 | 0.3 |

**Table 10.** Multivariate logistic regression NW.
* significant at P≤0.30

Patient feeling (p=0.206) and nurse feeling (p=0.195) were the only variables that remained in the model.

## PART 2AB. COMBINING MANUAL VITAL SIGN OBSERVATIONS AND NURSE WORRY INDICATORS.

Based on previous results, the variables that showed to be of significantly increase (P ≤0.30) the chance of experiencing an adverse event were used together in a multivariate logistic regression to make a combined end model.

| variables | β | SE | p-value* | Odds ratio | Normal scores | Ridge scores |
|---|---|---|---|---|---|---|
| Heart rate (beats/min) | 0.680 | 0.120 | 0.000 | 2.1 | 7 | 5 |
| Respiratory rate (breaths/min) | 0.166 | 0.100 | 0.095 | 2.0 | 2 | 1 |
| Temperature (°C) | -0.809 | 0.228 | 0.000 | 0.4 | 0 | 0 |
| SPO2 (%) | -0.216 | 0.087 | 0.013 | 0.8 | 0 | 0 |
| Systolic blood pressure (mmHg) | 0.701 | 0.257 | 0.006 | 2.0 | 7 | 6 |
| Nurse feeling | 0.735 | 0.666 | 0.270 | 2.0 | 7 | 6 |
| Worry temperature | 0.908 | 0.658 | 0.168 | 2.5 | 9 | 7 |
| Patient feeling | 0.856 | 0.621 | 0.169 | 2.4 | 9 | 7 |

**Table 15.** The combined end model
* significant at P<0.30

All the manual vital sign observations were shown to be significantly increasing the chance of an adverse. Therefore all variables should be used in the model. For the nurse worry indicators, worry about breathing, temperature, patient feeling and nurse feeling were initially used based on the univariate and multivariate analysis of the previous chapter. After backward selection only worry about temperature, nurse feeling, and patient feeling made it into the end model (table 15).

Based on the beta coefficient of this end model, the new continuous warning score (CWS), is developed (table 16). SPO2 and Temperature still show negative values indicating that they decrease the chance of an adverse event when present. For this reason, these variable receive zero points. The ranges of the vital signs are identical to the MEWS; only the scores they receive are based on the beta coefficients of the new model. A heart rate value that would normally score one point on the MEWS will according to the CWS score seven points. Heart rate values between 111-129 or ≤40 will normally score 2 points

on the MEWS. On CWS they will score 7x2, 14 points. Besides the 'normal' logistic regression, ridge regression was also performed with the same variables. The scores are obtained similar to the normal logistic regression. The CWS with scores based on the ridge regression will be called the CWS2. (table 17)

| | 0 | 2 | 4 | 6 | 7 | 9 | 14 | 21 |
|---|---|---|---|---|---|---|---|---|
| Heart rate (BMP) | 51-100 | | | | 101-110<br>41-50 | | 111-129<br>≤40 | ≥130 |
| Respiratory rate | 9-14 | 15-20 | 21-29<br>≤8 | ≥30 | | | | |
| Systolic blood pressure | 101-200 | | | | 81-100 | | ≥201<br>71-80 | ≤70 |
| Worry nurse feeling | | | | | Present | | | |
| Worry about temperature | | | | | | Present | | |
| Worry patient feeling | | | | | | Present | | |

**Table 16.** Continuous warning score (CWS)

| | 0 | 1 | 2 | 3 | 5 | 6 | 7 | 10 | 12 | 15 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Heart rate | 51-100 | | | | 101-110<br>41-50 | | | 111-129<br>≤40 | | ≥130 | |
| Respiratory rate | 9-14 | 15-20 | 21-29<br>≤8 | ≥30 | | | | | | | |
| Systolic blood pressure | 101-200 | | | | | 81-100 | | | ≥201<br>71-80 | | ≤70 |
| Worry nurse feeling | | | | | | Present | | | | | |
| Worry about temperature | | | | | | | Present | | | | |
| Worry patient feeling | | | | | | | Present | | | | |

**Table 17.** Continuous warning score 2 (based on ridge regression) (CWS2)

## PART 2C. MEWS ON CONTINUOUS DATA (CMEWS)

The regular MEWS was tested on the continuous data to create baseline measures of the continuous data to compare with the newly developed models. Different time intervals are compared. In current practice, a MEWS of ≥3, will set off a rapid response. On the continuous data, a MEWS of ≥3 on every time interval gave a sensitivity of 100% and a specificity of 0%. Higher thresholds scores gave a better sensitivity, specificity. In the table, the score that gave the highest sensitivity, specificity, PV+, and PV- are shown. (table 18)(figure 7).

| Time interval (min) | optimal threshold | sensitivity (%) | Specificity (%) | PV+ (%) | PV- (%) | AUC |
|---|---|---|---|---|---|---|
| **5** | 10 | 62 | 57 | 29 | 84, | 0.67 |
| **10** | 10 | 54 | 60 | 27 | 82 | 0.66 |
| **30** | 9 | 69 | 47 | 27 | 84 | 0.67 |
| **60** | 9 | 62 | 66 | 34 | 86 | 0.72 |
| **120** | 9 | 62 | 77 | 43 | 88 | 0.80 |
| **240** | 9 | 62 | 87 | 58 | 89 | 0.78 |
| **480** | 8 | 77 | 72 | 44 | 79 | 0.81 |

**Table 18. 11** Outcomes MEWS on continuous data

The time intervals from 5 to 30 minutes showed AUC values that suggest the MEWS on continuous data for these time intervals have poor discriminative power, which means that at these intervals, the MEWS cannot adequately distinguish between true negative and true positives. At 60 minutes or more, the AUC is above 0.7, which means that the discriminative power is acceptable. The best performing time interval is at 480 minutes, with the highest combination of sensitivity/specificity at a threshold value of 8 points.

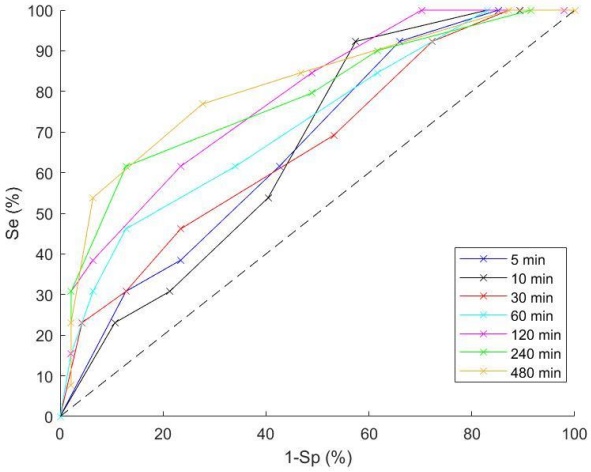



**Figure 7.** ROC CMEWS at different time intervals

### Time to detect

Time to detect for the CMEWS on continuous data was calculated. The threshold used to define detection with the CMEWS was a CMEWS of ≤3. Time to detect was calculated for 5-60-480 minutes.

Five minutes showed an average detection time 89.6 hours (3.7 days). 60 and 480 minute time intervals gave averages time to detect of 87.2 hours (3.6 days) and 78.2(3.3 days) respectively. These results show that continuous data with five-minute interval show that adverse events are detected a day earlier than when vital signs are manually obtained. There is not much differences between time intervals.

### Missed events

With a threshold of ≥3 no events were missed. However there were also a lot of false positives. The thresholds that had the best combination of sensitivity and specificity(table 18) were also analyzed at 5-60-480 minute interval to see what kind events were missed.

At five minute interval, the type events that were missed were pneumonia (5x), wound leakage, urinary tract infection and myocardial infarct. At the 60 and 480 minute time interval, missed events were similar as 5 minute only pneumonia missed was 4 and 3 times, respectively.

There was one patient with who was classified in the control group but had the highest MEWS scores of all the patients. This patient, however, had experienced two events, but they occurred before the start of the experiment, which means these events were not included in the analysis. The patient experienced a pleural emphysema followed by an anastomotic leak which can explain the high MEWS score.

## PART 2C. CONTINOUS WARNING SCORE ON CONTINUOUS DATA

### *CWS*

The new model that was developed, based on the manual observations and nurse worry indicators, was tested on the continuous data. The running averaged for different time intervals was used. Sensitivity, specificity, PV+, and PV- were calculated for every possible score of the new scoring system to chose the best rapid response trigger.

Patients could score a minimum of 0 points and a maximum of 52 points on the CWS based on their vital signs. There was not much difference between 5 and 10 min, 30 and 60 min and 120-240 min. Based on the AUC, the new model performed significantly better on every time interval compared to the traditional MEWS on manual obtained and CMEWS. The AUC for every time interval is above 0.7, which shows that the CWS has acceptable discriminative powers.

| Time interval (min) | Optimal threshold | sensitivity (%) | Specificity (%) | PV+ (%) | PV- (%) | AUC |
|---|---|---|---|---|---|---|
| **5** | 27 | 92 | 49 | 34 | 96 | 0.75 |
| **60** | 27 | 77 | 72 | 44 | 92 | 0.77 |
| **480** | 20 | 77 | 74 | 46 | 92 | 0.86 |

**Table 12. Outcomes CWS on continuous data**

At the 5 minute interval, the specificity is the same as the MEWS on manual vital sign data and even lower than the specificity of nurse worry alone (table 21). The time intervals of 60 minutes and more show significant improvements compared to the traditional MEWS and NW in sensitivity and specificity. The larges time step of 480 min is showing the highest AUC of 0.86, with a sensitivity and specificity of 77 % and 74%, respectively (table 19).

### Time to detect

The average time to detect an adverse event with the CWS in the time interval of 5 minutes was 104.7 hours (4.3 days) for a score of 27. The average time to detect for the 60-minute interval at a score of 27 was 80 hours (3.3) days. The 480-minute interval gave an average time to detect of 67 hours (2.8 days). These results show that the CWS is detecting adverse event earlier than the normal MEWS on manually obtained data and continuous data.

### Missed events

The adverse events that were missed by the CWS when using the optimal threshold points are analyzed for 5-60-480 minutes. At five minute interval, the types of events that were missed were atrial fibrillation (3x), pneumonia(5x) and urinary tract infection. 60 and 480-minute interval had the same amount and type of adverse events missed. The types of events that were missed were pneumonia (2x) and urinary tract infection. Similar as in the regular MEWS on continuous data, the patient who experienced a pleural emphysema and anastomotic leak before being included in the study was falsely classified as positive at every score.

### CWS 2

The CWS 2 was based on the ridge regression. With the CWS2 patient could score a minimum of 0 and maximum of 38 points based on vital signs. The results of sensitivity, specificity, PW+, PW- and AUC were similar to the normal CWS. Time to detect at 5-60-480 minute intervals and the missed adverse event at the optimal thresholds were the same as the normal CWS. Results for all the time intervals are found in Appendix(C-D).

## OVERALL SUMMERY CWS VS MEWS

Overall, the CWS/CWS 2 performed better than the normal MEWS on manual data and on the continuous dataset (CMEWS) at every time interval. When using the CMEWS of ≥3 as threshold, results were worse than manual observations. The most optimal score for the CMEWS was ≥9. However even the most optimal score performed worse than the CWS/CWS2.

Choosing the best time interval for the CWS depends on the trade-off between earlier detection and detecting more events. On average the running average with a 5-minute averaging interval detects adverse events a day earlier than when using a 60 minute interval and one and a half-day earlier than the 480 minute interval. However, using the 60 and 480 minute intervals more events are detected.

When using the CWS on continuous data, the running average of the 60-minute interval with a threshold of 27 seems to be the best option. Sensitivity and specificity with 60 minute interval is significantly better than 5 minute interval and almost the same as a 480 minute interval. The AUC of the 480 minute interval is best; however, the average detection time is close to the average time of detection of the current MEWS on the manually obtained dataset. The 60 minutes interval detects, on average, a day earlier than the 480 minute interval.

The overview of the 60-minute CWS model compared to normal MEWS and nurse Worry on manually obtained data and the traditional MEWS on continuous data at 60-minute interval is shown below (table 20)(figure 8)

| | Sensitivity (%) | Specificity (%) | PV+ (%) | PV- (%) | AUC |
|---|---|---|---|---|---|
| CWS/CWS 2 | 77 | 72 | 44 | 92 | 0.77 |
| CMEWS (≥3) | 100 | 0 | 22 | 0 | 0.72 |
| CMEWS (≥9) | 62 | 66 | 34 | 86 | 0.72 |
| MEWS ≥3 | 64 | 54 | 28 | 84 | 0.63 |
| Nurse Worry | 100 | 49 | 36 | 100 | 0.54 |

**Table 20.** Summary table of all the models

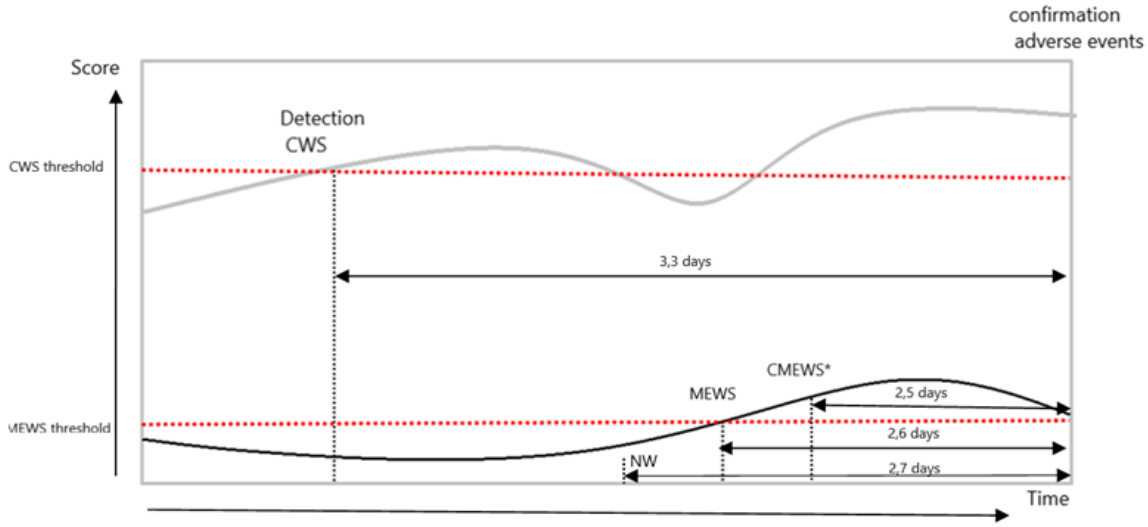



**Figure 8.** Time to detect for all the models. Greyline = CWS, blackline = MEWS

*CMEWS ≥9

# DISCUSSION

The results show that the CWS model significantly outperforms the NW and the traditional MEWS in both the manual and the continuous dataset (CMEWS). The CWS can detect more adverse events and can even detect them on average a day earlier than the MEWS. Detecting deterioration earlier is critical in preventing major complication that can result in impairment, function loss, longer days spend in hospital and even mortality (2-5). Several other studies had already emphasized the importance of nurse intuition (16,18-19). Comparing the CWS results to those of the traditional MEWS on continuous data, confirms that adding nurse worry indicators to the model is very valuable in better and earlier detection of adverse events.

MEWS scores based on manual vital signs observations and nurse worry were analyzed to see how well they performed in predicting adverse events. The results of this analysis showed a sensitivity of 100% and a specificity 49% with an AUC 0.6 for MEWS. Literature shows similar results with sensitivities between 70-90% and specificities between 48-80% with an AUC 0.68-0,8 (9,49,50). Sensitivity and specificity are often inversely proportional to one another, which explains the combination of high sensitivity and the lower specificity. The AUC of 0,6 shows that MEWS has poor discriminative power in this dataset. Nurse worry showed lesser performances compared to the MEWS, even though specificity in nurse worry is slighter better with 54%. The AUC was 0.54 which shows that nurse worry is very bad in discriminating between event and no event, and suggest that for the investigated dataset it is mostly random chance whether patients are classified in the right event or no event group. However, results of a study with a bigger sample size shows an AUC of 0.87 for nurse worry, suggesting that nurse worry is indeed a good predictor (19).

The most important variables of the MEWS were found to be heart rate, respiratory rate and systolic blood pressure. These results are comparable to literature (44,49-50). The results of the analysis suggest that SPo2 and temperature are not good predictors for adverse events. The multivariate analysis showed that they had significant negative beta coefficients, suggesting that they decrease the chance of adverse events. Due to their significance, temperature and SPo2 were used in the end model but not scored in the CWS.

Remarkably, literature is showing that SPo2 and temperature are very important predictors (42,9,44). Upon closer inspection, this also becomes apparent by the type of events that are missed by CWS with a threshold of 27. The events that are classified as false negatives, when using a threshold of 27, were events such as pneumonia and urinary tract infections. These complications are usually accompanied by body temperature rises, whereas other vital signs might in the beginning still be normal. Moreover, pneumonia causes oxygen saturation to drop and therefore be a critical measure in detection of pneumonia (45). One of the reasons that SPo2 and temperature showed negative values in the analyses could be explained by the lack of data of these measures (40).

The most important nurse worry indicators were worry about temperature, patient feeling and nurse feeling. A lot of indicators were not included in the analysis due to small number of filled in nurse worry checklists . However, a study showed way more indicators to be significantly important in predicting adverse events (19).

The most important variables were placed into the end model. Unfortunately, the dataset contained a relatively small number of events. To reduce bias and overfitting in construction of a model, best practice is to have at least ten events per variable. This means that the number of events should be at least 60 events in the dataset for a model that uses six variables. However, the dataset that was used only had 13 events. Ridge regression was used to overcome this problem; yet, the results were similar compared to the outcomes of the standard logistic regression. One advance of the ridge method is that the potential scores are more closely spaced and thus cause less substantial differences in possible scores. This can make it easier to interpret for nurses.

The traditional MEWS was tested on the continuous data to make baseline measures (CMEWS). With a threshold of ≥3 the model gave a sensitivity of 100% and a specificity of 0% which suggest that all the patients had a MEWS of ≥3, at some point during continuous monitoring. When increasing the threshold, specificity increased. The optimal threshold score was observed between 8 and 10, depending on time interval. Nevertheless, optimal scores were still lower than the CWS model. This shows that using the traditional MEWS on the

continuous data is undesirable and highlights the importance of nurse worry indicators.

The time interval of 8 hours with CWS gave the largest AUC (0.86) with a sensitivity of 77% and a specificity of 74%. PV+ and PV- were 46% and 92% respectively. However, if a running average of 8 hours is used, then the first 8 hours of data would not be scored, which may result in missed deterioration that can occur within these 8 hours. Furthermore, taking the running average of 8 hours does not provide any faster detection relative to the traditional MEWS scoring system. A time interval of 8 hours had a similar time to detect compared to the traditional MEWS in current practice . A running average of 60 min is therefore preferred. The 60 min interval with a threshold of 27 points shows a slightly lower AUC (0.77) but similar sensitivity (77 %) ,specificity (70%), PV+ (42%) and PV- (92%), meaning that the same amount of events are detected. However, the events are, on average, detected a day earlier than using manually obtained vital signs. With this smaller interval, the benefits of continuous data in earlier detection can thus more optimally be used. PV+ and PV- should be interpreted with some caution as these are based on the prevalence of the disease. However prevalence is not known, so for this thesis it is assumed to be equal to the number of events registered in the dataset.

Noticeable in the continuous data is that, on average, respiratory rate is much higher than in the manual observations. The average from continuous measurements is 26 breaths/min patients with an event and 24 breaths/min in patients without, whereas the manual obtained averages are 17 breaths/min and 16 breaths/min, respectively. These values highlight a difference between the method used by nurses to count the breathing frequency and the frequency that is registered using sensors. This discrepancy in respiratory rate between manually obtained observations and sensor data is found in more studies (20,41). One reason for these differences could be because nurses often consider counting breathing frequency a waste of time and will often fill in an estimate, which results in underestimating respiratory rate (41-42). The big difference also suggests that the ranges of what is considered a ‘normal rate’ and what is considered a deviating rate need reassessment for continuous monitoring. In the CWS model, the scales of the normal MEWS were used, which causes a lot of patients to receive points for deviating respiratory rate. Only 8% of the patients

scored zero points on the CWS, which again suggest that the ranges of categorized vital signs need to be reassessed for the continuous data when used for detecting adverse events.

### Limitation and strengths

There are some limitations to this thesis. First of all, secondary data where used in a retrospective design. As a result there was no control over the way the data was collected and missing data was unavoidable. Last observation carried forward was used to impute the missing data. This imputation of the data can influence the predictive models by underestimating or overestimating results. Underestimating the importance of temperature and Spo2 might be a direct result of this approximation.

The small dataset and small number of events also influences sensitivity/specificity/PV+/PV-/AUC. With these small numbers in the calculations the right or wrong categorization of one patient can already change these diagnostic values significantly.

Moreover, in defining adverse events and detection some assumptions were made. In the definition of detection it is assumed that when the threshold is reached for the first time, the event is detected. However one cannot be certain that the high score is related to the event. This assumption about detection also influences detecting time. Also, it is assumed that the patients without event and no Nurse Worry or MEWS were true negatives. However, it could be that the events are not yet detected by the standard protocol but are nevertheless present. Furthermore, the inclusion of adverse events might influence the results. Patients can have had an adverse event which was not included but does result in deviating vital signs. This means that sensitivity and specificity should be interpreted with some caution due to the possibility that false positives might be true positives; while, the events were not yet detected by standard protocol. This also means that there is a possibility that some true negatives are false negatives.

Another limitation is that systolic pressure and nurse worry were taken from manual observations and included in the CWS/CWS2/CMEWS as they could not be measured continuously. At the moment in time that a nurse worry was

expressed, or systolic pressure was deviating outside the normal range, the scores of these measures were added to the total scores in the following four hours. This introduces some uncertainty in the results as these added scores might to not represent the correct condition of the patient in the following four hours. The calculated score might therefore be underestimating or overestimating the actual condition of the patient.

Furthermore, due to the relatively small datasets, some of the nurse worry indicators had very few observations and were taken out of the analysis. However, a study that looked into these nurse worry indicators showed that most of the indicators are good predictors of events (19). Also, the limited amount of data could have negatively influenced the outcomes of the diagnostic values of nurse worry in the manual observations

A key strength of the study is the use of several different types of methods to investigate the predictive powers of NW and MEWS. Furthermore combining NW indicators and MEWS and the following investigation on continuous data had not been done before and is showing promising results. Also, the traditional MEWS was tested on the continuous data to serve as a baseline for comparison of results on continuous data.

Another strength is the investigation into different time intervals for running average scoring values. This analysis showed that the choice of time interval influenced the results significantly, underlining the importance of selecting a suitable time interval for CWS monitoring.

## Practical implications

Results show that the introduction of continuous monitoring has the potential to support earlier and better detection of adverse events. This increases patient safety on general wards. Not only will this be beneficial for patients, but also for the nurses and physicians. Literature showed that nurses play a crucial role in detecting deterioration and triggering of the rapid response system (12). Unfortunately nurses often encounter high workloads, fear of lacking skills or knowledge to provide quantifiable information to set-off a proper response. Furthermore they experience scarcity of equipment and difficulty in

communication between nurses. Introducing continuous monitoring has high potential to resolve these problems. It will reduce workload and provide nurses with constant information about their patients' status. This constant source of information can encourage nurses to be more proactive, have more knowledge about how the patient is doing and empowers them to be in control. The results of this work not only show the benefits of continuous monitoring but also emphasize the importance of nurse worry. This is also complementary to the literature were the importance of nurse intuition was already described (18-19) and was therefore should be a validated criteria to set off rapid response.

Unfortunately, if the CWS is introduced, NW indicators cannot be measured continuously. They should be conducted by nurses every nurse shift or more in case they worry about specific patients. The scores taken from these worry indicators should be automatically added to the score of the vital signs. When nurse worry indicators are activated, the score should automatically be added to the score of the vitals sign for the following four hours. If the worry indicators are not activated again after four hours, the NW scores are not added to the vital signs anymore until they are reactivated.

## Recommendations

Results of this study shows that the introduction of continuous monitoring with the CWS yields great potential and are therefore recommended to be used in practice. Unfortunately there was limited amount data which makes the results uncertain. More research should be conducted focusing on the nurse indicators in combination with the vital signs to find the most optimal combination. Also research should be focusing on patterns in deteriorating vital signs. These patterns should be used to trigger an alarm and could possibly be used to suggest likely diagnoses. Some of the events are missed because they are not deviating "enough" to set off an alarm but are very important signs of deterioration.

Due to the small data set all the data before the events was analyzed. It is recommended to use the 48 hours before an event. Literature showed that physiological deterioration can be detected up to 48 hours before an event. By only including the 48 hours before an event in the analysis, the chance that the

deviating vital signs are related to the event is bigger. It is also recommended to reinvestigate suitable vital signs ranges for continuous vital signs scoring. The difference in mean measured respiratory rate between manual and continuous obtained data was already discussed. However, other vital signs also showed differences in mean between manual and continuous observations. For example, the temperature in the continuous data is measures based on auxiliary skin temperature, whereas for the temperature in manually obtained data tympanic thermometers were used. The average temperature in the continuous dataset and the manual data set of all the patients were 36,2 and 37,1, respectively. This confirms that the temperature measured from the auxiliary skin is lower and therefore, ranges of the current MEWS are not representative for this type of temperature measurement. The mean of Spo2 in the continuous data was also slightly lower than in the manually obtained vital signs with 92% and 95%, respectively. Additional research should be conducted to determine ranges that are fitting continuous vital signs, based on the sensory equipment that is selected.

Furthermore, wireless automatic blood pressure was not used in this study when continuously monitoring patients. However, systolic blood pressure is showing to be a significant predictor for adverse events. Therefore, it is also recommended to monitor blood pressure continuously. Isansys wearable sensors were used for heart rate, respiratory rate, SPo2 and temperature. Isansys also supplies wireless blood pressure sensors that can be connected to the digitalization station.

Finally, besides the absolute threshold, an alarm that shows when vital signs are slowly changing might be a valuable asset for continuous monitoring. This relative score can, for example, be the difference between the previous score and the current score. When this number is positive for an extended period of time, it suggests that the score is increasing and vital signs are worsening. Nurses can use this to pay extra attention to patients and are alerted of potential deterioration before the threshold is even reached.

# CONCLUSION

Introducing continuous monitoring has great potential in the earlier and more accurate detection of adverse events. The inclusion of nurse worry indicators in the prediction model results in an increased predictive performance of the model. This is complementary to literature, were the importance of nurse worry was already touched upon by some studies (18-19). Respiratory rate, heart rate and systolic blood pressure showed to be the biggest predictors of adverse events in the manual obtained vital signs observations. Worry about temperature, nurse- and patient worry were the most important nurse indicators. These variables combined performed significantly better on continuous data than the traditional MEWS on manual and on continuous data. Introducing continuous monitoring on the surgical wards and using vital signs and nurse worry indicators to detect deterioration in patient condition both show great potential towards earlier and better detection of adverse event.  Better and earlier detection of adverse event can results in more patient safety, lower mortality rates, reduce workload of nurses and better communication between health professionals.

Future research should focus on establishing more certainty about scoring ranges for the vital signs in continuous monitoring, pattern recognition in vital signs in combination with nurse worry indicators and look into the added value of relative scores besides absolute scores to indicate deterioration.

# LITERATURE


1. Cypress, B. S. (2013). Transfer Out of Intensive Care. *Dimensions of Critical Care Nursing*, *32*(5), 244–261. https://doi.org/10.1097/DCC.0b013e3182a07646
2. Duus, C. L., Aasvang, E. K., Olsen, R. M., Sørensen, H. B. D., Jørgensen, L. N., Achiam, M. P., & Meyhoff, C. S. (2018). Continuous vital sign monitoring after major abdominal surgery-Quantification of micro events. *Acta Anaesthesiologica Scandinavica*, *62*(9), 1200–1208. https://doi.org/10.1111/aas.13173
3. Escobar, G. J., Laguardia, J. C., Turk, B. J., Ragins, A., Kipnis, P., & Draper, D. (2012). Early detection of impending physiologic deterioration among patients who are not in intensive care: Development of predictive models using data from an automated electronic medical record. *Journal of Hospital Medicine*, *7*(5), 388–395. https://doi.org/10.1002/jhm.1929
4. Young, M. P., Gooder, V. J., McBride, K., James, B., & Fisher, E. S. (2003). Inpatient transfers to the intensive care unit: Delays are associated with increased mortality and morbidity. *Journal of General Internal Medicine*, *18*(2), 77–83. https://doi.org/10.1046/j.1525-1497.2003.20441.x
5. Odell, M. (2015). Detection and management of the deteriorating ward patient: an evaluation of nursing practice. *Journal of Clinical Nursing*, *24*(1–2), 173–182. https://doi.org/10.1111/jocn.12655
6. Stolldorf, D. P., & Jones, C. B. (2015). Deployment of rapid response teams by 31 hospitals in a statewide collaborative. *Joint Commission Journal on Quality and Patient Safety*, *41*(4), 186–192. https://doi.org/10.1016/s1553-7250(15)41024-4
7. Morgan, R., Lloyd-Williams, F., Wright, M., & Morgan-Warren, R. (1997). *An early warning scoring system for detecting developing critical illness*.
8. Franklin, C., & Mathew, J. (1994). Developing strategies to prevent inhospital cardiac arrest: Analysing responses of physicians and nurses in the hours before the event. *Critical Care Medicine*, *22*(2), 244–247. https://doi.org/10.1097/00003246-199402000-00014
9. Gardner-Thorpe, J., Love, N., Wrightson, J., Walsh, S., & Keeling, N. (2006). The value of Modified Early Warning Score (MEWS) in surgical in-patients: A prospective observational study. *Annals of the Royal College of Surgeons of England*, *88*(6), 571–575. https://doi.org/10.1308/003588406X130615
10. Petersen, J. A., Mackel, R., Antonsen, K., & Rasmussen, L. S. (2014). Serious adverse events in a hospital using early warning score - What went wrong? *Resuscitation*, *85*(12), 1699–1703. https://doi.org/10.1016/j.resuscitation.2014.08.037

11. Wood, C., Chaboyer, W., & Carr, P. (2019). How do nurses use early warning scoring systems to detect and act on patient deterioration to ensure patient safety? A scoping review. In *International Journal of Nursing Studies*, 94, 166–178. Elsevier Ltd. https://doi.org/10.1016/j.ijnurstu.2019.03.012
12. Rao, A. D., & DeVita, M. (2011). RRS's General Principles. In *Textbook of Rapid Response Systems* (pp. 13–18). Springer New York. https://doi.org/10.1007/978-0-387-92853-1_2
13. Olsen, S. L., Søreide, E., Hillman, K., & Hansen, B. S. (2019). Succeeding with rapid response systems – a never-ending process: A systematic review of how health-care professionals perceive facilitators and barriers within the limbs of the RRS. In *Resuscitation*, 144, 75–90. Elsevier Ireland Ltd. https://doi.org/10.1016/j.resuscitation.2019.08.034
14. Trinkle, R. M., & Flabouris, A. (2011). Documenting Rapid Response System afferent limb failure and associated patient outcomes. *Resuscitation*, *82*(7), 810–814. https://doi.org/10.1016/j.resuscitation.2011.03.019
15. Ludikhuize, J., Smorenburg, S. M., De Rooij, S. E., & De Jonge, E. (2012). *Identification of deteriorating patients on general wards; measurement of vital parameters and potential effectiveness of the Modified Early Warning Score*. Journal of criterical care, 27(4), 424-424. https://doi.org/10.1016/j.jcrc.2012.01.003
16. Kim, W. Y., Shin, Y. J., Lee, J. M., Huh, J. W., Koh, Y., Lim, C. M., & Hong, S. B. (2015). Modified early warning score changes prior to cardiac arrest in general wards. *PLoS ONE*, *10*(6). https://doi.org/10.1371/journal.pone.0130523
17. McGain, F., Cretikos, M. A., Jones, D., van Dyk, S., Buist, M. D., Opdam, H., Pellegrino, V., Robertson, M. S., & Bellomo, R. (2008). Documentation of clinical review and vital signs after major surgery. In *Medical Journal of Australia*, 168(7) ,380–383. https://doi.org/10.5694/j.1326-5377.2008.tb02083.x
18. Romero-Brufau, S., Gaines, K., Nicolas, C. T., Johnson, M. G., Hickman, J., & Huddleston, J. M. (2019). *The fifth vital sign? Nurse worry predicts inpatient deterioration within 24 hours*. JAMIA open, 2(4), 465-470. 10.1093/jamiaopen/ooz033
19. Douw, G., Huisman-de Waal, G., van Zanten, A. R. H., van der Hoeven, J. G., & Schoonhoven, L. (2016). Nurses' "worry" as predictor of deteriorating surgical ward patients: A prospective cohort study of the Dutch-Early-Nurse-Worry-Indicator-Score. *International Journal of Nursing Studies*, 59, 134-140. 10.1016/j.ijnurstu.2016.04.006
20. Weenk, M., Koeneman, M., Van De Belt, T. H., Engelen, L. J. L. P. G., Van Goor, H., & Bredie, S. J. H. (2019). *Wireless and continuous monitoring of vital signs in patients at the general*

*ward*. Resuscitation, 136, 47-53. https://doi.org/10.1016/j.resuscitation.2019.01.017
21. Jones, D., Rubulotta, F., & Welch, J. (2016). Rapid response teams improve outcomes: yes. In *Intensive care medicine* , 42 (4) 593–595. Springer. https://doi.org/10.1007/s00134-016-4219-5
22. Neale, G., Woloshynowych, M., & Vincent, C. (2001). Exploring the causes of adverse events in NHS hospital practice. In *Journal of the Royal Society of Medicine*, 94(7), 322–330. Royal Society of Medicine Press Ltd. https://doi.org/10.1177/014107680109400702
23. Maharaj, R., Raffaele, I., & Wendon, J. (2015). Rapid response systems: a systematic review and meta-analysis. *Critical Care*, *19*(1), 254. https://doi.org/10.1186/s13054-015-0973-y
24. Ludikhuize, J., A. Hamming, J. E. de, and B. G. Fikkers. 2011. Rapid response systems in The Netherlands. Jt.Comm J.Qual.Patient Saf 37:138-44, 97.
25. Keene, C. M., Kong, V. Y., Clarke, D. L., & Brysiewicz, P. (2017). The effect of the quality of vital sign recording on clinical decision making in a regional acute care. *Chinese Journal of Traumatology, 20(5), 283-287.* 10.1016/j.cjtee.2016.11.008
26. Stenhouse, C., Coates, S., Tivey, M., Allsop, P., & Parker, T. (2000). *Prospective evaluation of a modified Early Warning Score to aid earlier detection of patients developing critical illness on a general surgical ward*. Proceedings of the intensive care Society. https://doi.org/10.1093/bja/84.5.663
27. Odell, M., Victor, C., & Oliver, D. (2009). Nurses' role in detecting deterioration in ward patients: systematic literature review. *Journal of Advanced Nursing*, *65*(10), 1992–2006. https://doi.org/10.1111/j.1365-2648.2009.05109.x
28. Santiano, N., Young, L., Hillman, K., Parr, M., Jayasinghe, S., Baramy, L. S., Stevenson, J., Heath, T., Chan, C., Claire, M., & Hanger, G. (2009). Analysis of Medical Emergency Team calls comparing subjective to "objective" call criteria. *Resuscitation*, *80*(1), 44–49. https://doi.org/10.1016/j.resuscitation.2008.08.010
29. Gao, H., McDonnell, A., Harrison, D.A., Moore, T., Adam, S., Daly, K., Esmonde, L., Goldhill, D.R., Parry, G.J., Rashidian, A., Subbe, C.P., Harvey, S., 2007. Systematic review and evaluation of physiological track and trigger warning systems for identifying at-risk patients on the ward. Intens. Care Med. 33 (4), 667–679.
30. Isansys (2020).*Wearable sensors, Lifetouch.* Retrieved from: https://www.isansys.com/en/Wearable-Sensors
31. Insansys (2020). Wearable sensors, LifeTemp. Retrieved from: https://www.isansys.com/en/Wearable-Sensors
32. Downey, C. L., Chapman, S., Randell, R., Brown, J. M., & Jayne, D. G. (2018). The impact of continuous versus

intermittent vital signs monitoring in hospitals: A systematic review and narrative synthesis. In *International Journal of Nursing Studies,* 84, 19–27). Elsevier Ltd. https://doi.org/10.1016/j.ijnurstu.2018.04.013
33. Daabiss, M. (2011). American society of anaesthesiologists physical status classification. In *Indian Journal of Anaesthesia*, 55(2), 111–115. Wolters Kluwer -- Medknow Publications. https://doi.org/10.4103/0019-5049.79879
34. Twisk, J.W.R (2016). *Inleiding in de toegepaste biostatistiek.* Bohn stafleu vaan Loghun, Houten.
35. Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression.pdf* (pp. 161–164). https://doi.org/10.1038/461726a
36. Noordhuizen, J.P.T.M., Frankena, K., Thrsusfield, M.V., & Graat, E.A.M. (2017). *Application of quantitative methods in veterinary epidemiology.* Wageningen Academic Publishers. The Netherlands.
37. Mehta, H. B., Mehta, V., Girman, C. J., Adhikari, D., & Johnson, M. L. (2016). Regression coefficient–based scoring system should be used to assign weights to the risk index. *Journal of Clinical Epidemiology*, *79*, 22–28. https://doi.org/10.1016/j.jclinepi.2016.03.031
38. Peduzzi, P., Concato, J., Feinstein, A. R., & Holford, T. R. (1995). Importance of events per independent variable in proportional hazards regression analysis II. Accuracy and precision of regression estimates. *Journal of Clinical Epidemiology*, *48*(12), 1503–1510. https://doi.org/10.1016/0895-4356(95)00048-8
39. Pavlou, M., Ambler, G., Seaman, S. R., Guttmann, O., Elliott, P., King, M., & Omar, R. Z. (2015). How to develop a more accurate risk prediction model when there are few events. *BMJ (Online)*, *351*. https://doi.org/10.1136/bmj.h3868
40. *A gentle introduction to logistic regression and lasso regularisation using R | Eight to Late*. (n.d.). Retrieved from https://eight2late.wordpress.com/2017/07/11/a-gentle-introduction-to-logistic-regression-and-lasso-regularisation-using-r/
41. Boatin, A. A., Wylie, B. J., Goldfarb, I., Azevedo, R., Pittel, E., Ng, C., & Haberer, J. E. (2016). Wireless Vital Sign Monitoring in Pregnant Women: A Functionality and Acceptability Study. *Telemedicine and E-Health*, *22*(7), 564–571. https://doi.org/10.1089/tmj.2015.0173
42. Flenady, T., Dwyer, T., & Applegarth, J. (2017). Accurate respiratory rates count: So should you! In *Australasian Emergency Nursing Journa, 20(1)*, 45–47. Elsevier Ltd. https://doi.org/10.1016/j.aenj.2016.12.003
43. Solomon, R. S., Corwin, G. S., Barclay, D. C., Quddusi, S. F., & Dannenberg, M. D. (2016). Effectiveness of rapid response

teams on rates of in-hospital cardiopulmonary arrest and mortality: A systematic review and meta-analysis. In *Journal of Hospital Medicine* 11(6), 438–445. John Wiley and Sons Inc. https://doi.org/10.1002/jhm.2554

44. Churpek, M. M., Adhikari, R., & Edelson, D. P. (2016). The value of vital sign trends for detecting clinical deterioration on the wards. *Resuscitation*, *102*, 1–5. https://doi.org/10.1016/j.resuscitation.2016.02.005
45. Majumdar, S. R., Eurich, D. T., Gamble, J. M., Senthilselvan, A., & Marrie, T. J. (2011). Oxygen saturations less than 92% are associated with major adverse events in outpatients with pneumonia: A population-based cohort study. *Clinical Infectious Diseases*, *52*(3), 325–331. https://doi.org/10.1093/cid/ciq076
46. Joshi, M., Ashrafian, H., Aufegger, L., Khan, S., Arora, S., Cooke, G., & Darzi, A. (2018). *Wearable sensors to improve detection of patient deterioration*. *Expert Review of Medical Devices , 16(2),145-154.* https://doi.org/10.1080/17434440.2019.1563480
47. Brown H, Terrence J, Vasquez P (2014). Continuous monitoring in an inpatient medical-surgical unit: a controlled clinical trial. Am J Med, 127(11),226-232. https://doi.org/10.1016/j.amjmed.2013.12.004
48. Breteler, M. J. M., Huizinga, E., Van Loon, K., Leenen, L. P. H., Dohmen, D. A. J., Kalkman, C. J., & Blokhuis, T. J. (2018). Reliability of wireless monitoring using a wearable patch sensor in high-risk surgical patients at a step-down unit in the Netherlands: A clinical validation study. *BMJ Open*, *8*(2). https://doi.org/10.1136/bmjopen-2017-020162
49. Cuthbertson, B. H., Boroujerdi, M., McKie, L., Aucott, L., & Prescott, G. (2007). Can physiological variables and early warning scoring systems allow early recognition of the deteriorating surgical patient? *Critical Care Medicine*, *35*(2), 402–409. https://doi.org/10.1097/01.CCM.0000254826.10520.87
50. Suwanpasu, S., & Sattayasomboon, Y. (2016). Accuracy of Modified Early Warning Scores for Predicting Mortality in Hospital: A Systematic Review and Meta-analysis. *IMedPub Journals*, *2*(2), 20. https://doi.org/10.21767/2471-8505.100029

# Niet-pluis checklist

**Invullen:**
Elke dienst: Bij de eerste controles
*indien geen controles aan begin van de dienst; invullen bij 1e ronde langs patiënt*

Extra: Bij een toenemend niet-pluis gevoel

1. Algemeen

| Datum: | Tijd van invullen: ……:…… h |
|---|---|
| Naam van patiënt | |

2. Hoeveel jaar bent u werkzaam als verpleegkundige op deze afdeling?
- ☐ Ik ben nog in opleiding
- ☐ 0-1 jaar
- ☐ 1-2 jaar
- ☐ 2-5 jaar
- ☐ > 5 jaar

3. Bent u bekend met deze patiënt?
- ☐ Ja — *Ik heb in de afgelopen 7 dagen één of meerdere dagen voor deze patiënt gezorgd als verantwoordelijk verpleegkundige*
- ☐ Deels — *Ik ken de patiënt, maar ben niet de verantwoordelijk verpleegkundige geweest de afgelopen 7 dagen*
- ☐ Nee — *Het is een (zo goed als) nieuwe patiënt voor mij*

4. Bent u ongerust over de huidige status van de patiënt?
- ☐ Ik ben niet ongerust → *Ga naar vraag 5*
- ☐ Ik ben een klein beetje ongerust
- ☐ Ik ben redelijk ongerust
- ☐ Ik ben zeer ongerust

4b. Zo ja, hoe laat begon deze ongerustheid?

……:…… h

4c. Bij ongerustheid: wat is de onderliggende reden?*
*Zie achterzijde voor toelichting*
- ☐ Verandering in ademhaling
- ☐ Verandering in circulatie
- ☐ Temperatuur
- ☐ Mentale verandering
- ☐ Agitatie
- ☐ Pijn
- ☐ Overige, *namelijk ...*
- ☐ Onverwacht beloop
- ☐ Patiënt geeft aan zich niet goed te voelen
- ☐ Familie geeft aan dat het niet goed zit
- ☐ Verpleegkundig gevoel

5. Is er aanleiding om actie te ondernemen of de patiënt extra in de gaten te houden?*
- ☐ Nee → *Ga naar vraag 6*
- ☐ Ja, op basis van de recente vitale functies of MEWS
- ☐ Ja, op basis van hoe de patiënt zich voelt
- ☐ Ja, op basis van mijn eigen ongerustheid
- ☐ Ja, op basis van andere zaken, *namelijk...*

5b. Zo ja, welke actie is of wordt in gang gezet?*
- ☐ Extra oogje houden op de patiënt
- ☐ Extra controles van de vitale functies
- ☐ Bellen/overleg met arts
- ☐ Oproepen interventie team (SIT)
- ☐ Anders, *namelijk....*

6. Opmerkingen/toelichting:

*Meerdere antwoorden mogelijk

Formulier inleveren in MoViSign brievenbus, bedankt!

## *Toelichting indicatoren Niet-pluis gevoel*

***Vraag 4C***

1. *Verandering in ademhaling*
   - *Hoorbare ademhaling*
   - *Kortademig*
   - *Niet in volzinnen kunnen praten*
   - *Gebruik hulpademhalingsspieren*

2. *Verandering in circulatie*
   - *Verandering kleur: bleek/grauw*
   - *Transpireren/klam*
   - *Koud aanvoelen*
   - *Verminderde doorbloeding*
   - *Oedemen*

3. *Temperatuur*
   - *Rillingen*

4. *Mentale verandering*
   - *Apathie/slaperig*
   - *Verward*

5. *Agitatie*
   - *Rusteloos*
   - *Angstig*

6. *Pijn*
   - *Nieuwe pijn*
   - *Toenemende of aanhoudende pijn*

7. *Een niet verwacht beloop*
   - *Geen vooruitgang*
   - *Opgezette buik/misselijk/braken*
   - *Bloeding*
   - *Duizelig*
   - *Flauw vallen*
   - *Vallen*

8. *Patiënt geeft aan zich niet goed te voelen*
   - *Niet goed te voelen*
   - *Gevoel van naderend onheil*

9. *Familie geeft aan dat het niet goed zit*
   - *Verandering in gedrag*
   - *Verandering in houding*
   - *Ziet er niet goed uit*
   - *Blik in de ogen*

10. *Verpleegkundig gevoel*
    - *Verandering in gedrag*
    - *Verandering in houding*
    - *Ziet er niet goed uit*
    - *Blik in de ogen*

*Indicatoren zijn gebaseerd op de **Dutch Early Nurse Worry Indicator Score (DENWIS)**, ontwikkeld in het RadboudUMC.*

***Vragen?** Neem contact op met één van de MoViSign ambassadeurs of onderzoekers (088 708 4515 of movisign@zgt.nl)*

# APPENDIX B

**Patient characteristics**

- *Age:* Numerical data.
- *Gender:* Gender of the patient. Categorical data; 'Male' or 'Female'.
- *Weight:* Weight of the patient. Numerical data.
- *Height:* Height of the patient. Numerical data.
- *ASA score:* Classification of health status (American Society of anesthesiologists). A subjective score that used by anesthesiologists). To indicate the overall health of the patient preoperative. ASA contains five classes (I-V) (34).
    - I. ' patient is a completely healthy fit patient.'
    - II. 'Patient has mild systemic disease'.
    - III. 'Patient has a severe systemic disease that is not incapacitating'.
    - IV. 'Patient has an incapacitating disease that is a constant threat to life.'
    - V. 'A dying patient who is not expected to live 24 hours with or without surgery.'
- *Type of surgery:* Indicator for the hospital admission. Either elective oesophagal- or gastric surgery or traumatic hip fracture surgery (elderly of 70+). Categorical data.
- *Pre-existing medical condition*: Presence of specific comorbidities preoperative. Categorical data.
    - Cardiac
    - Vascular
    - Diabetes
    - Pulmonary
    - Neurologic/psychiatric
    - Gastrointestinal
    - Urogenital
    - Thrombotic
    - Neuromuscular

- Infection diseases
- Endocrine
- Other

- *Adverse event:* An adverse event is defined as an undesirable experience occurring during the study. An adverse event is included when it meets a set of criteria. First, it is a postoperative complication classified as Clavien Dindo class of II (table 4), diagnosed according to standard protocol. Furthermore, the complication should have started at least 24 hours after the start of the measurement period. Categorical data.
- *Complication timing:* The time that adverse event was detected and confirmed according to standard protocol. Date and time data.
- *Type of adverse event:* The type of complication that is detected and confirmed according to standard protocol. Text data.

# APPENDIX C

| Time interval (min) | Optimal threshold | sensitivity (%) | Specificity (%) | PV+ (%) | PV- (%) | AUC |
|---|---|---|---|---|---|---|
| **5** | 27 | 92 | 49 | 34 | 96 | 0.75 |
| **10** | 27 | 85 | 53 | 34 | 92 | 0.73 |
| **30** | 27 | 85 | 66 | 41 | 94 | 0.78 |
| **60** | 27 | 77 | 72 | 44 | 92 | 0.77 |
| **120** | 27 | 77 | 78 | 53 | 89 | 0.81 |
| **240** | 22 | 69 | 81 | 50 | 90 | 0.81 |
| **480** | 20 | 77 | 74 | 46 | 92 | 0.86 |

**Table 13. OUtcomes CWS all time intervals**

| Time interval | score | Sensitivity (%) | Specificity(%) | PV+ (%) | PV- (%) | AUC |
|---|---|---|---|---|---|---|
| **5** | 20 | 54 | 77 | 39 | 85 | 0.75 |
| **10** | 19 | 85 | 53 | 34 | 92 | 0.72 |
| **30** | 18 | 85 | 64 | 40 | 94 | 0.78 |
| **60** | 18 | 77 | 70 | 42 | 92 | 0.77 |
| **120** | 18 | 77 | 77 | 48 | 92 | 0.80 |
| **240** | 17 | 69 | 81 | 50 | 90 | 0.78 |
| **480** | 14 | 77 | 74 | 46 | 92 | 0.85 |

**Table 22.** Outcomes cws2 all time intervals

# APPENDIX D

## CWS

***5 min continue***

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 2 | 100 | 0 | 22 | 100 |
| 4 | 100 | 0 | 22 | 100 |
| 6 | 100 | 0 | 22 | 100 |
| 7 | 100 | 0 | 22 | 100 |
| 9 | 100 | 8.51 | 23.56 | 100 |
| 11 | 100 | 14.89 | 24.89 | 100 |
| 13 | 100 | 17.02 | 25.37 | 100 |
| 14 | 100 | 17.02 | 25.37 | 100 |
| 15 | 100 | 17.02 | 25.37 | 100 |
| 16 | 100 | 21.28 | 26.38 | 100 |
| 18 | 92.31 | 21.28 | 24.85 | 90.75 |
| 10 | 23.08 | 97.87 | 75.35 | 78.00 |
| 11 | 7.69 | 97.87 | 50.45 | 0.00 |
| 12 | 0 | 100 | 0.00 | 0.00 |
| 13 | 0 | 100 | 0.00 | 0.00 |
| 20 | 92.31 | 25.53 | 25.91 | 92.17 |
| 21 | 92.31 | 25.53 | 25.91 | 92.17 |
| 22 | 92.31 | 44.68 | 32.00 | 95.37 |
| 25 | 92.31 | 46.81 | 32.86 | 95.57 |
| 27 | 92.31 | 48.94 | 33.77 | 95.76 |
| 29 | 30.31 | 91.49 | 50.11 | 82.32 |
| 36 | 15.38 | 97.87 | 67.07 | 80.39 |
| 43 | 7.69 | 97.87 | 50.45 | 78.99 |
| 52 | 0 | 97.877 | 0.00 | 77.63 |

**10 min**

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 2 | 100 | 0 | 22 | 100 |
| 4 | 100 | 0 | 22 | 100 |
| 6 | 100 | 2.13 | 22.37 | 100 |
| 7 | 100 | 2.13 | 22.37 | 100 |
| 9 | 100 | 10.64 | 23.99 | 100 |
| 11 | 100 | 19.15 | 25.86 | 100 |
| 13 | 100 | 21.28 | 26.38 | 100 |
| 14 | 100 | 21.28 | 26.38 | 100 |
| 15 | 92.31 | 21.28 | 24.85 | 90.75 |

| | | | | |
|---|---|---|---|---|
| 16 | 92.31 | 23.4 | 25.37 | 91.52 |
| 18 | 92.31 | 23.4 | 25.37 | 91.52 |
| 20 | 92.31 | 27.66 | 26.47 | 92.73 |
| 21 | 92.31 | 27.66 | 26.47 | 92.73 |
| 22 | 84.62 | 48.94 | 31.85 | 91.86 |
| 25 | 84.62 | 51.06 | 32.78 | 92.17 |
| 27 | 84.62 | 53.19 | 33.77 | 92.46 |
| 29 | 30.77 | 91.49 | 50.49 | 82.41 |
| 36 | 15.38 | 97.89 | 67.28 | 80.40 |
| 43 | 7.69 | 97.87 | 50.45 | 78.99 |
| 52 | 0 | 97.87 | 0.00 | 77.63 |

**30 min**

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 2 | 100 | 0 | 22 | 100 |
| 4 | 100 | 0 | 22 | 100 |
| 6 | 100 | 8.5 | 23.56 | 100 |
| 7 | 100 | 8.5 | 23.56 | 100 |
| 9 | 100 | 21.28 | 26.38 | 100 |
| 11 | 100 | 27.66 | 28.05 | 100 |
| 13 | 100 | 29.78 | 28.66 | 100 |
| 14 | 100 | 29.78 | 28.66 | 100 |
| 15 | 92.31 | 31.91 | 27.66 | 93.64 |
| 16 | 92.31 | 36.17 | 28.97 | 94.34 |
| 18 | 92.31 | 36.17 | 28.97 | 94.34 |
| 20 | 84.62 | 40.43 | 28.60 | 90.31 |
| 21 | 84.62 | 40.43 | 28.60 | 90.31 |
| 22 | 84.62 | 63.83 | 39.75 | 93.64 |
| 25 | 84.62 | 63.83 | 39.75 | 93.64 |
| 27 | 84.62 | 65.95 | 41.21 | 93.83 |
| 29 | 30.77 | 93.87 | 58.61 | 82.78 |
| 36 | 15.39 | 97.87 | 67.08 | 80.40 |
| 43 | 7.68 | 97.87 | 50.42 | 78.99 |
| 52 | 0 | 97.87 | 0.00 | 77.63 |

**60 min**

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 2 | 100 | 0 | 22 | 100 |
| 4 | 100 | 0 | 22 | 100 |
| 6 | 100 | 12.77 | 24.43 | 100 |
| 7 | 100 | 12.77 | 24.43 | 100 |

| | | | | |
|---|---|---|---|---|
| 9 | 100 | 27.66 | 28.05 | 100 |
| 11 | 100 | 36.17 | 30.65 | 100 |
| 13 | 100 | 40.42 | 32.13 | 100 |
| 14 | 100 | 40.42 | 32.13 | 100 |
| 15 | 76.92 | 44.68 | 28.17 | 87.28 |
| 16 | 76.92 | 48.94 | 29.82 | 88.26 |
| 18 | 76.92 | 48.94 | 29.82 | 88.26 |
| 20 | 76.92 | 53.19 | 31.67 | 89.10 |
| 21 | 76.92 | 53.19 | 31.67 | 89.10 |
| 22 | 76.92 | 70.21 | 42.14 | 91.51 |
| 25 | 76.92 | 70.21 | 42.14 | 91.51 |
| 27 | 76.92 | 72.34 | 43.96 | 91.74 |
| 29 | 30.92 | 93.62 | 57.75 | 82.77 |
| 36 | 15.38 | 97.87 | 67.07 | 80.39 |
| 43 | 7.69 | 97.87 | 50.45 | 78.99 |
| 52 | 0 | 97.87 | 0.00 | 77.63 |

**120 min**

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 2 | 100 | 0 | 22 | 100 |
| 4 | 100 | 0 | 22 | 100 |
| 6 | 100 | 21.28 | 26.38 | 100 |
| 7 | 100 | 21.28 | 26.38 | 100 |
| 9 | 100 | 27.66 | 28.05 | 100 |
| 11 | 100 | 36.17 | 30.65 | 100 |
| 13 | 100 | 46.8 | 34.65 | 100 |
| 14 | 100 | 46.8 | 34.65 | 100 |
| 15 | 76.92 | 46.8 | 28.97 | 87.79 |
| 16 | 76.92 | 53.19 | 31.67 | 89.10 |
| 18 | 76.92 | 53.19 | 31.67 | 89.10 |
| 20 | 76.92 | 55.32 | 32.69 | 89.47 |
| 21 | 76.92 | 55.32 | 32.69 | 89.47 |
| 22 | 76.92 | 76.6 | 48.11 | 92.17 |
| 25 | 76.92 | 76.6 | 48.11 | 92.17 |
| 27 | 76.92 | 78.72 | 50.48 | 92.36 |
| 29 | 30.76 | 95.74 | 67.07 | 83.06 |
| 36 | 15.38 | 97.87 | 67.07 | 80.39 |
| 43 | 7.69 | 97.87 | 50.45 | 78.99 |
| 45 | 0 | 97.87 | 0.00 | 77.63 |

**240 min**

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 2 | 100 | 0 | 22 | 100 |
| 4 | 100 | 0 | 22 | 100 |
| 6 | 100 | 25.53 | 27.47 | 100 |
| 7 | 100 | 25.53 | 27.47 | 100 |
| 9 | 100 | 34.04 | 29.95 | 100 |
| 11 | 100 | 40.42 | 32.13 | 100 |
| 13 | 100 | 51.06 | 36.56 | 100 |
| 14 | 100 | 51.06 | 36.56 | 100 |
| 15 | 76.92 | 53.19 | 31.67 | 89.10 |
| 16 | 76.92 | 57.59 | 33.84 | 89.84 |
| 18 | 76.92 | 61.7 | 36.16 | 90.46 |
| 20 | 76.92 | 63.83 | 37.49 | 90.75 |
| 21 | 76.92 | 63.83 | 37.49 | 90.75 |
| 22 | 69.23 | 80.85 | 50.49 | 90.31 |
| 25 | 61.53 | 80.85 | 47.54 | 88.17 |
| 27 | 61.53 | 82.9 | 50.37 | 88.43 |
| 29 | 23.08 | 95.74 | 60.44 | 81.53 |
| 36 | 7.69 | 97.87 | 50.45 | 78.99 |
| 43 | 7.69 | 97.87 | 50.45 | 78.99 |
| 52 | 0 | 0 | 0.00 | 0.00 |

**480 min**

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 2 | 100 | 0 | 22 | 100 |
| 4 | 100 | 0 | 22 | 100 |
| 6 | 100 | 34.04 | 29.95 | 100 |
| 7 | 100 | 34.04 | 29.95 | 100 |
| 9 | 100 | 44.55 | 33.72 | 100 |
| 10 | 100 | 44.68 | 33.77 | 100 |
| 11 | 100 | 51.06 | 36.56 | 100 |
| 12 | 100 | 51.06 | 36.56 | 100 |
| 13 | 100 | 61.7 | 42.41 | 100 |
| 14 | 100 | 61.7 | 42.41 | 100 |
| 15 | 76.92 | 65.96 | 38.93 | 91.02 |
| 16 | 76.92 | 68.96 | 41.14 | 91.37 |

| | | | | |
|---|---|---|---|---|
| 18 | 76.92 | 70.21 | 42.14 | 91.51 |
| 20 | 76.92 | 74.47 | 45.94 | 91.96 |
| 21 | 76.92 | 74.47 | 45.94 | 91.96 |
| 23 | 61.54 | 89.36 | 62.00 | 89.17 |
| 24 | 53.84 | 89.36 | 58.80 | 87.28 |
| 25 | 53.85 | 89.36 | 58.81 | 87.29 |
| 27 | 53.85 | 89.36 | 58.81 | 87.29 |
| 29 | 23.08 | 97.87 | 75.35 | 81.85 |
| 33 | 7.69 | 97.87 | 50.45 | 78.99 |

**5min**

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 1 | 100 | 0 | 22 | 100 |
| 2 | 100 | 0 | 22 | 100 |
| 3 | 100 | 0 | 22 | 100 |
| 5 | 100 | 8.51 | 23.56 | 100 |
| 6 | 100 | 8.51 | 23.56 | 100 |
| 7 | 100 | 8.51 | 23.56 | 100 |
| 8 | 100 | 14.89 | 24.89 | 100 |
| 9 | 100 | 14.89 | 24.89 | 100 |
| 10 | 100 | 17.02 | 25.37 | 100 |
| 11 | 100 | 17.02 | 25.37 | 100 |
| 12 | 92.31 | 21.28 | 24.85 | 90.75 |
| 13 | 92.31 | 21.28 | 24.85 | 90.75 |
| 14 | 92.31 | 25.53 | 25.91 | 92.17 |
| 15 | 92.31 | 44.68 | 32.00 | 95.37 |
| 16 | 92.31 | 44.68 | 32.00 | 95.37 |
| 18 | 92.31 | 46.81 | 32.86 | 95.57 |
| 19 | 92.31 | 48.94 | 33.77 | 95.76 |
| 20 | 53.85 | 76.6 | 39.36 | 85.48 |
| 21 | 30.77 | 89.36 | 44.92 | 82.07 |
| 22 | 23.08 | 95.74 | 60.44 | 81.53 |
| 26 | 15.38 | 97.87 | 67.07 | 80.39 |
| 31 | 7.69 | 97.87 | 50.45 | 78.99 |
| 38 | 0 | 97.87 | 0.00 | 77.63 |

**10 min**

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 1 | 100 | 0 | 22 | 100 |
| 2 | 100 | 0 | 22 | 100 |
| 3 | 100 | 0 | 22 | 100 |
| 4 | 100 | 2.13 | 22.37 | 100 |
| 5 | 100 | 10.64 | 23.99 | 100 |
| 6 | 100 | 10.64 | 23.99 | 100 |
| 7 | 100 | 10.64 | 23.99 | 100 |
| 8 | 100 | 19.15 | 25.86 | 100 |
| 9 | 100 | 19.15 | 25.86 | 100 |
| 10 | 100 | 21.28 | 26.38 | 100 |
| 11 | 92.31 | 21.28 | 24.85 | 90.75 |

| | | | | |
|---|---|---|---|---|
| 12 | 92.31 | 23.4 | 25.37 | 91.52 |
| 14 | 92.31 | 27.66 | 26.47 | 92.73 |
| 15 | 84.61 | 48.94 | 31.85 | 91.85 |
| 16 | 84.61 | 48.94 | 31.85 | 91.85 |
| 18 | 84.62 | 51.19 | 32.84 | 92.19 |
| 19 | 84.62 | 53.19 | 33.77 | 92.46 |
| 20 | 46.15 | 78.72 | 37.95 | 83.83 |
| 21 | 30.77 | 89.36 | 44.92 | 82.07 |
| 22 | 23.08 | 95.74 | 60.44 | 81.53 |
| 26 | 15.39 | 97.87 | 67.08 | 80.40 |
| 31 | 7.69 | 97.87 | 50.45 | 78.99 |
| 38 | 0 | 97.87 | 0.00 | 77.63 |

**30 min**

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 1 | 100 | 0 | 22 | 100 |
| 2 | 100 | 0 | 22 | 100 |
| 3 | 100 | 0 | 22 | 100 |
| 4 | 100 | 8.51 | 23.56 | 100 |
| 5 | 100 | 21.28 | 26.38 | 100 |
| 6 | 100 | 21.28 | 26.38 | 100 |
| 7 | 100 | 21.28 | 26.38 | 100 |
| 8 | 100 | 25.53 | 27.47 | 100 |
| 9 | 100 | 27.66 | 28.05 | 100 |
| 10 | 100 | 29.79 | 28.66 | 100 |
| 11 | 92.31 | 31.91 | 27.66 | 93.64 |
| 12 | 92.31 | 36.17 | 28.97 | 94.34 |
| 13 | 92.31 | 36.17 | 28.97 | 94.34 |
| 14 | 84.61 | 40.42 | 28.60 | 90.30 |
| 15 | 84.61 | 61.7 | 38.39 | 93.43 |
| 16 | 84.61 | 61.7 | 38.39 | 93.43 |
| 18 | 84.61 | 63.83 | 39.75 | 93.63 |
| 19 | 84.61 | 63.83 | 39.75 | 93.63 |
| 20 | 46.15 | 87.23 | 50.48 | 85.17 |
| 21 | 30.77 | 81.49 | 31.92 | 80.67 |
| 22 | 23.08 | 95.74 | 60.44 | 81.53 |
| 26 | 15.38 | 97.87 | 67.07 | 80.39 |
| 31 | 7.69 | 97.87 | 50.45 | 78.99 |
| 38 | 0 | 97.87 | 0.00 | 77.63 |

**60 min**

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 1 | 100 | 0 | 22 | 100 |
| 2 | 100 | 0 | 22 | 100 |
| 3 | 100 | 0 | 22 | 100 |
| 4 | 100 | 12.76 | 24.43 | 100 |
| 5 | 100 | 27.66 | 28.05 | 100 |
| 6 | 100 | 27.66 | 28.05 | 100 |
| 7 | 100 | 27.66 | 28.05 | 100 |
| 8 | 100 | 31.91 | 29.29 | 100 |
| 9 | 100 | 36.17 | 30.65 | 100 |
| 10 | 100 | 40.42 | 32.13 | 100 |
| 11 | 76.92 | 44.68 | 28.17 | 87.28 |
| 12 | 76.92 | 48.94 | 29.82 | 88.26 |
| 13 | 76.92 | 48.94 | 29.82 | 88.26 |
| 14 | 76.92 | 53.19 | 31.67 | 89.10 |
| 15 | 76.92 | 68.09 | 40.47 | 91.27 |
| 16 | 76.92 | 68.09 | 40.47 | 91.27 |
| 18 | 76.92 | 70.21 | 42.14 | 91.51 |
| 19 | 76.92 | 70.21 | 42.14 | 91.51 |
| 20 | 38.46 | 87.23 | 45.93 | 83.40 |
| 21 | 30 | 91.48 | 49.83 | 82.25 |
| 22 | 23.08 | 95.74 | 60.44 | 81.53 |
| 23 | 23.08 | 97.87 | 75.35 | 81.85 |
| 26 | 15.38 | 97.87 | 67.07 | 80.39 |
| 31 | 7.69 | 97.87 | 50.45 | 78.99 |
| 38 | 0 | 97.87 | 0.00 | 77.63 |

**120 min**

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 1 | 100 | 0 | 22 | 100 |
| 2 | 100 | 0 | 22 | 100 |
| 3 | 100 | 0 | 22 | 100 |
| 4 | 100 | 21.27 | 26.38 | 100 |
| 5 | 100 | 27.66 | 28.05 | 100 |
| 6 | 100 | 27.66 | 28.05 | 100 |
| 7 | 100 | 27.66 | 28.05 | 100 |
| 8 | 100 | 31.91 | 29.29 | 100 |
| 9 | 100 | 36.17 | 30.65 | 100 |

| | | | | |
|---|---|---|---|---|
| 10 | 100 | 46.81 | 34.65 | 100 |
| 11 | 76.92 | 46.81 | 28.97 | 87.79 |
| 12 | 76.92 | 53.19 | 31.67 | 89.10 |
| 13 | 76.92 | 53.19 | 31.67 | 89.10 |
| 14 | 76.92 | 55.32 | 32.69 | 89.47 |
| 15 | 76.92 | 74.47 | 45.94 | 91.96 |
| 18 | 76.92 | 76.6 | 48.11 | 92.17 |
| 19 | 76.92 | 76.6 | 48.11 | 92.17 |
| 20 | 38.46 | 91.49 | 56.04 | 84.05 |
| 21 | 30.77 | 93.62 | 57.63 | 82.74 |
| 22 | 23.08 | 95.74 | 60.44 | 81.53 |
| | 23.08 | 97.87 | 75.35 | 81.85 |
| 26 | 15.39 | 97.87 | 67.08 | 80.40 |
| 31 | 7.69 | 97.87 | 50.45 | 78.99 |
| 33 | 0 | 97.87 | 0.00 | 77.63 |

**240 hours**

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 1 | 100 | 0 | 22 | 100 |
| 2 | 100 | 0 | 22 | 100 |
| 3 | 100 | 0 | 22 | 100 |
| 4 | 100 | 25.53 | 27.47 | 100 |
| 5 | 100 | 34.04 | 29.95 | 100 |
| 6 | 100 | 34.04 | 29.95 | 100 |
| 7 | 100 | 34.04 | 29.95 | 100 |
| 8 | 100 | 36.17 | 30.65 | 100 |
| 9 | 100 | 40.43 | 32.13 | 100 |
| 10 | 100 | 53.06 | 36.56 | 100 |
| 11 | 76.92 | 53.19 | 31.67 | 89.10 |
| 12 | 76.92 | 59.58 | 34.93 | 90.15 |
| 13 | 76.92 | 61.7 | 36.16 | 90.46 |
| 14 | 76.92 | 63.83 | 37.49 | 90.75 |
| 15 | 69.23 | 78.72 | 47.85 | 90.07 |
| 17 | 69.23 | 80.85 | 50.49 | 90.31 |
| 18 | 61.53 | 80.85 | 47.54 | 88.17 |
| 19 | 61.54 | 80.85 | 47.54 | 88.17 |
| 20 | 30.76 | 91.49 | 50.48 | 82.41 |
| 21 | 23.07 | 93.62 | 50.49 | 81.18 |
| 22 | 15.38 | 95.74 | 50.45 | 80.05 |
| 26 | 7.69 | 97.87 | 50.45 | 78.99 |
| 31 | 0 | 0 | 0.00 | 0.00 |

**480 min**

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 1 | 100 | 0 | 22 | 100 |
| 2 | 100 | 0 | 22 | 100 |
| 3 | 100 | 0 | 22 | 100 |
| 4 | 100 | 34.04 | 29.95 | 100 |
| 6 | 100 | 42.55 | 32.93 | 100 |
| 7 | 100 | 42.55 | 32.93 | 100 |
| 8 | 100 | 46.8 | 34.65 | 100 |
| 9 | 100 | 51.06 | 36.56 | 100 |
| 10 | 100 | 61.7 | 42.41 | 100 |
| 11 | 76.92 | 65.96 | 38.93 | 91.02 |
| 12 | 76.92 | 68.09 | 40.47 | 91.27 |
| 13 | 76.92 | 70.21 | 42.14 | 91.51 |
| 14 | 76.92 | 74.47 | 45.94 | 91.96 |
| 15 | 61.53 | 87.23 | 57.61 | 88.94 |
| 17 | 61.53 | 89.36 | 61.99 | 89.17 |
| 18 | 53.85 | 89.36 | 58.81 | 87.29 |
| 19 | 53.85 | 89.36 | 58.81 | 87.29 |
| 20 | 30.08 | 93.62 | 57.08 | 82.60 |
| 21 | 23.08 | 95.74 | 60.44 | 81.53 |
| 22 | 15.38 | 97.87 | 67.07 | 80.39 |
| 26 | 7.69 | 97.87 | 50.45 | 78.99 |
| 33 | 0 | 97.87 | 0.00 | 77.63 |

**5 min**

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 1 | 100 | 0 | 22 | 100 |
| 2 | 100 | 0 | 22 | 100 |
| 3 | 100 | 0 | 22 | 100 |
| 4 | 100 | 0 | 22 | 100 |
| 5 | 100 | 0 | 22 | 100 |
| 6 | 100 | 0 | 22 | 100 |
| 7 | 100 | 10.64 | 23.99 | 100.00 |
| 8 | 100 | 14.89 | 24.89 | 100.00 |
| 9 | 92.31 | 34.04 | 28.30 | 94.01 |
| 10 | 61.54 | 57.45 | 28.97 | 84.12 |
| 11 | 38.46 | 76.59 | 31.66 | 81.52 |
| 12 | 30.77 | 87.23 | 40.46 | 81.71 |
| 13 | 0 | 100 | 0.00 | 78.00 |

**10 min**

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 1 | 100 | 0 | 22 | 100 |
| 2 | 100 | 0 | 22 | 100 |
| 3 | 100 | 0 | 22 | 100 |
| 4 | 100 | 0 | 22 | 100 |
| 5 | 100 | 0 | 22 | 100 |
| 6 | 100 | 0 | 22 | 100 |
| 7 | 100 | 10.64 | 23.99 | 100 |
| 8 | 100 | 17.02 | 25.37 | 100 |
| 9 | 92.31 | 42.55 | 31.19 | 95.15 |
| 10 | 53.85 | 59.57 | 27.31 | 82.07 |
| 11 | 30.79 | 78.72 | 28.98 | 80.13 |
| 12 | 23.08 | 89.36 | 37.96 | 80.46 |
| 13 | 0 | 100 | 0.00 | 78.00 |

**30 min**

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 1 | 100 | 0 | 22 | 100 |
| 2 | 100 | 0 | 22 | 100 |

| | | | | |
|---|---|---|---|---|
| 3 | 100 | 0 | 22 | 100 |
| 4 | 100 | 0 | 22 | 100 |
| 5 | 100 | 0 | 22 | 100 |
| 6 | 100 | 0 | 22 | 100 |
| 7 | 100 | 2.13 | 22.37 | 100 |
| 8 | 92.31 | 12.77 | 22.99 | 100 |
| 9 | 69.23 | 46.81 | 26.85 | 84.36 |
| 10 | 46.15 | 76.59 | 35.73 | 83.45 |
| 11 | 30.77 | 87.23 | 40.46 | 81.71 |
| 12 | 23.08 | 95.74 | 60.44 | 81.53 |
| 13 | 0 | 100 | 0.00 | 78.00 |

**60 min**

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 1 | 100 | 0 | 22 | 100 |
| 2 | 100 | 0 | 22 | 100 |
| 3 | 100 | 0 | 22 | 100 |
| 4 | 100 | 0 | 22 | 100 |
| 5 | 100 | 0 | 22 | 100 |
| 6 | 100 | 2.13 | 22.37 | 100 |
| 7 | 100 | 17.02 | 25.37 | 100 |
| 8 | 84.61 | 38.3 | 27.89 | 100 |
| 9 | 61.54 | 65.96 | 33.77 | 85.88 |
| 10 | 46.15 | 87.23 | 50.48 | 85.17 |
| 11 | 30.77 | 93.62 | 57.63 | 82.74 |
| 12 | 15.38 | 97.87 | 67.07 | 80.39 |
| 13 | 0 | 100 | 0.00 | 78.00 |

**120 min**

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 1 | 100 | 0 | 22 | 100 |
| 2 | 100 | 0 | 22 | 100 |
| 3 | 100 | 0 | 22 | 100 |
| 4 | 100 | 0 | 22 | 100 |
| 5 | 100 | 0 | 22 | 100 |
| 6 | 100 | 2.13 | 22.37 | 100 |

| | | | | |
|---|---|---|---|---|
| 7 | 100 | 29.79 | 28.66 | 100 |
| 8 | 84.61 | 51.06 | 32.78 | 100 |
| 9 | 61.54 | 76.6 | 42.59 | 87.60 |
| 10 | 38.46 | 93.62 | 62.97 | 84.36 |
| 11 | 30.77 | 97.87 | 80.29 | 83.37 |
| 12 | 15.38 | 97.87 | 67.07 | 80.39 |
| 13 | 0 | 100 | 0.00 | 78.00 |

**240min**

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 1 | 100 | 0 | 22 | 100 |
| 2 | 100 | 0 | 22 | 100 |
| 3 | 100 | 0 | 22 | 100 |
| 4 | 100 | 0 | 22 | 100 |
| 5 | 100 | 0 | 22 | 100 |
| 6 | 100 | 8.5 | 23.56 | 100 |
| 7 | 90.08 | 38.3 | 29.17 | 100 |
| 8 | 79.61 | 51.06 | 31.45 | 100 |
| 9 | 61.53 | 87.23 | 57.61 | 88.94 |
| 10 | 30.77 | 97.87 | 80.29 | 83.37 |
| 11 | 23.07 | 97.87 | 75.34 | 81.85 |
| 12 | 7.67 | 97.87 | 50.39 | 78.98 |
| 13 | 0 | 100 | 0.00 | 78.00 |

**480 min**

| Score | Sensitivity | Specificity | PW+ | PW- |
|---|---|---|---|---|
| 1 | 100 | 0 | 22 | 100 |
| 2 | 100 | 0 | 22 | 100 |
| 3 | 100 | 0 | 22 | 100 |
| 4 | 100 | 0 | 22 | 100 |
| 5 | 100 | 0 | 22 | 100 |
| 6 | 100 | 12.77 | 24.43 | 87.79 |
| 7 | 84.61 | 53.19 | 33.77 | 81.85 |
| 8 | 76.92 | 72.34 | 43.96 | 78.99 |
| 9 | 53.85 | 93.62 | 70.42 | 78.00 |