Afnan S. Sharourou et al. Saudi J Er Med. 2023;4(1):001-005
https://doi.org/10.24911/SJEMed/72-1673204949

ORIGINAL ARTICLE

If the machine says "Normal" on the ECG, can I put the patient in the waiting room?: a study on the reliability of ECG software classification

Afnan S. Sharourou[1], Omar T. Idriss[1], Ahmed L. AlManzalawi[2], Amro M. Tayeb[3], Thamer A. AlGhamdi[4], Anas Fouad Hamam[5,6*]

ABSTRACT

Background: Electrocardiograms (ECGs) are one of the most basic and fundamental screening tools used in the emergency department (ED). Previous studies have shown machine diagnosis of ECG to be unreliable, the ECG machine does provide a simpler classification of: 1- Normal, 2- Otherwise normal, 3- Borderline, and 4- Abnormal printed on the ECG. We aim to investigate if machine classification could be used reliably as a screening test for triage.

Methods: This cross-sectional study was conducted from 1 to 14 June 2019 of ED at King Abdullah Medical Complex using electronic medical records. The ECGs were put into sets of 25 traces/set and then presented to 21 board-certified emergency medicine attending physicians (EMPs) to assess and decide on one of the actions: Put in the waiting area, see in triage, or admit immediately. The responses were analyzed for inter-subject correlation coefficient kappa (*k*).

Results: Of the 3,149 patients, 452 had ECGs done from which 200 ECGs were chosen at random. The inter-subject correlation coefficient was found to be 0.315 ± 0.187 denoting a fair to moderate correlation. From the ECG traces classified as "Normal" by the ECG machine, only 46% ended up in the waiting room. While almost 15% were admitted immediately to an ED bed. In contrast, 27% of those labeled as "Abnormal" ended up in the waiting room, while 44% were admitted to an ED bed.

Conclusion: The machine classification of the ECG traces unfortunately failed remarkably to predict the EMP's decision. As such, the assessment of the attending EMP remains a necessary and essential part of the assessment.

Keywords: Electrocardiography, triage, emergency service, hospital.

Introduction

Cardiac conditions are one of the most common causes of the presentation of the patient to an emergency department (ED). For this reason, electrocardiogram (ECG) is one of the most common tools used in the ED. It is a non-invasive, point-of-care investigation used to quickly screen for cardiac abnormalities such as ischemia or arrhythmia. In most EDs around the world, ECGs are usually done during triage. ECG is part of clinical pathway, or institution protocol, in accordance with the American Heart Association guidelines to minimize delays to patient care presenting with serious cardiac pathology [1].

Previous studies have shown that during a busy shift interruptions of the attending physician in the ED is associated with a negative detrimental effect on patient

Correspondence to: Anas Fouad Hamam
*Consultant of Emergency Medicine, King Fahad Armed Forces Hospital, Jeddah, Saudi Arabia.
Email: dr_anas_md@hotmail.com
Full list of author information is available at the end of the article.
Received: 08 January 2023 | **Accepted:** 13 March 2023

care [2,3]. Blocker et al. [4], found that, in an average shift, a general attending emergency physician is interrupted 11.21 times per hour. Which amounts to an interruption rate of once every 5.3 minutes. Thus, concentrating on one's clinical work, and maintaining situational focus, to carry out critical decisions becomes in itself a consuming and exhausting task. One which requires a honed mental discipline, which emergency physicians spend years learning to master.

Since performing an ECG on patients is one of the nursing functions, it has been reported in many previous studies, that the ECG reading skills of the majority of nurses are unfortunately, deficient and below average [5-7]. As such, once an ECG is completed for any patient, it is now necessary for the triaging nurse to show this ECG to the attending physician. This is a problem simply because it represents another interruption for the emergency physician. The clinical decision now falls to the attending physician on what to do with the patient.

With the development of the computerized ECG software, several studies showed that computerized diagnoses of ECG traces are not reliable. Physician assessments and professional evaluation of each case still remain the standard of care for identifying serious medical conditions. This is of course understandable, as computerized assessments do not take into account the many obvious and hidden information that a professional can take in with a simple glance of the patient and using his/her clinical sense. Information, such as the age, sex, presenting complaint, patient demeanor, how they look, and vitals, are obviously missing from the machine diagnosis of the ECG. However, the modern ECG machine also gives each ECG trace an overall read that is simple. The ECG traces are given a simple classification as being "Normal" or "Abnormal," regardless of the specific diagnosis of each individual trace. To our knowledge, no study has ever investigated if this simple ECG classification can be used as a crude screening tool to identify patients with "Normal" ECGs from those that have actual pathology. Is this classification sensitive and reliable to the point where it mirrors the decision taken by the emergency physician? If so, using this classification could go a long way to relieving some of the attention interruptions that the emergency physician is bombarded with by allowing them to focus on the task that they are performing at any given time. Allowing the nurse to use the machine classification, to direct the patient flow rather than having to ask the busy physician.

Methods

This cross-sectional study was conducted using data collected from the electronic medical record (EMR) at the King Abdullah Medical Complex (KAMCJ) in Jeddah, Saudi Arabia, a tertiary healthcare facility that sees approximately 120,000 patients/annum. The data were extracted from 00:00 on June 1, 2019 through till 23:59 on June 14, 2019. We assessed two main domains: 1- The ECG done for these patients and 2- The one-line history of their main presenting complaint that the triage nurse recorded for these patients in the EMR. This one-line history included their presenting complaint, age, sex, their chronic conditions, etc. (as an example see Figure 1). An approval from the Institutional Review Board (IRB) of the Jeddah Ministry of Health (MOH) administration was received, approval number A01319, on March 12, 2022.

We collected all the ECGs done during the 2-week period for all the patients that were presented to the ED. With each ECG, the one-line history was noted, and the simple machine classification was noted. The ECG machine used at KAMCJ was a Philips PageWriter® TC50 cardiograph (picture and technical specifications for the ECG software used are shown in Figure 2; with software revision A.07.3.07 and application revision 4.00.31010). This ECG machine brand is the most popular, and the one used in many, if not all, of the MOH hospitals. The PageWriter® TC50 assigns one of four simple classifications to each ECG trace done, according to its software, these are: normal, otherwise normal, borderline, and abnormal (shown in the red circle Figure 1).

The ECG traces collected were all cropped to remove any, and all, machine text, data, interpretation, and classification from them. The traces were further assigned a study identification number, and the respective history information, recorded by the triage nurse was added to each ECG (Figure 1 shows an example of how the ECG traces were prepared). After all the ECG traces were modified and prepped in this manner, 200 ECGs were

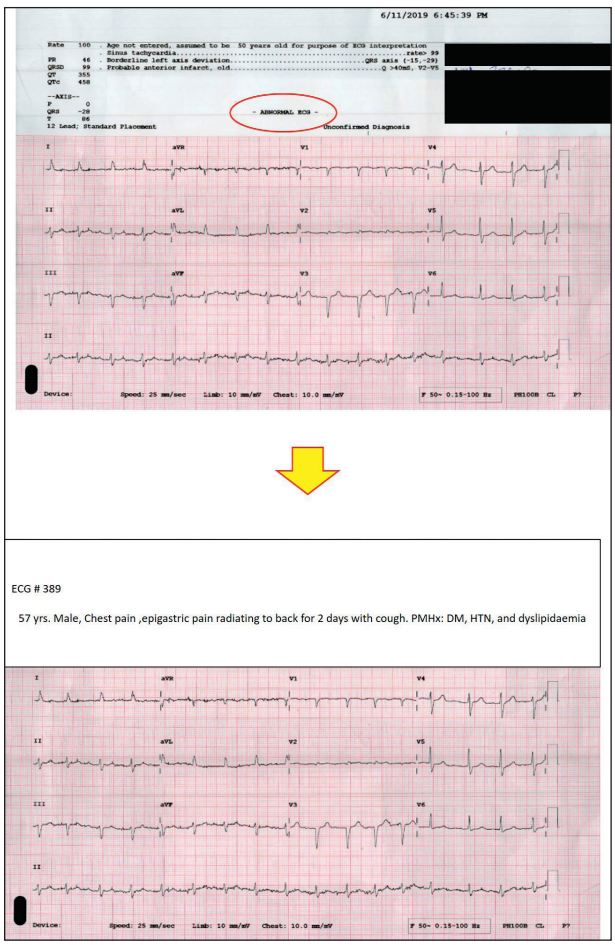
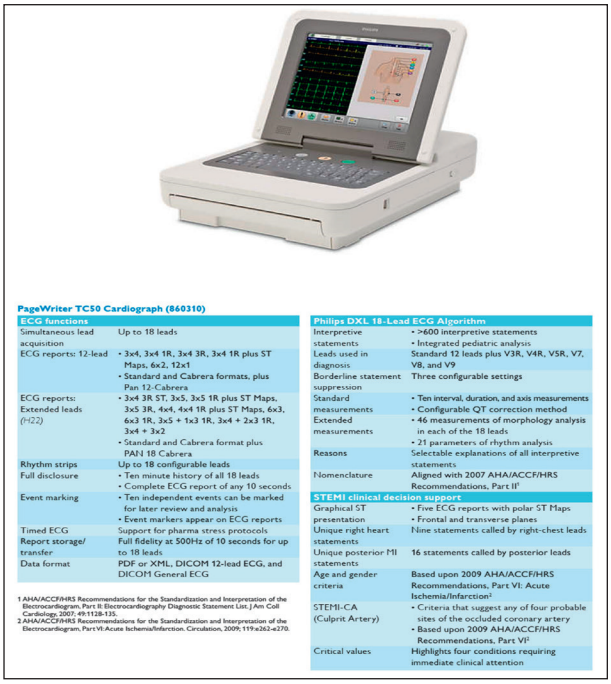


Figure 1. *How the ECG classifications have been masked and replaced with the presenting complaint.*



PageWriter TC50 Cardiograph (860310)

ECG functions	
Simultaneous lead acquisition	Up to 18 leads
ECG reports: 12-lead	• 3x4, 3x4 1R, 3x4 3R, 3x4 1R plus ST Maps, 6x2, 12x1 • Standard and Cabrera formats, plus Pan 12-Cabrera
ECG reports: Extended leads (H22)	• 3x4 3R ST, 3x5, 3x5 1R plus ST Maps, 3x5 3R, 4x4, 4x4 1R plus ST Maps, 6x3, 6x3 1R, 3x5 + 1x3 1R, 3x4 + 2x3 1R, 3x4 + 3x2 • Standard and Cabrera format plus PAN 18 Cabrera
Rhythm strips	Up to 18 configurable leads
Full disclosure	• Ten minute history of all 18 leads • Complete ECG report of any 10 seconds
Event marking	• Ten independent events can be marked for later review and analysis • Event markers appear on ECG reports
Timed ECG	Support for pharma stress protocols
Report storage/ transfer	Full fidelity at 500Hz of 10 seconds for up to 18 leads
Data format	PDF or XML, DICOM 12-lead ECG, and DICOM General ECG

1 AHA/ACCF/HRS Recommendations for the Standardization and Interpretation of the Electrocardiogram. Part II: Electrocardiography Diagnostic Statement List. J Am Coll Cardiology, 2007; 49:1128-135.
2 AHA/ACCF/HRS Recommendations for the Standardization and Interpretation of the Electrocardiogram. Part VI: Acute Ischemia/Infarction. Circulation, 2009; 119:e262-e270.

Philips DXL 18-Lead ECG Algorithm	
Interpretive statements	• >600 interpretive statements • Integrated pediatric analysis
Leads used in diagnosis	Standard 12 leads plus V3R, V4R, V5R, V7, V8, and V9
Borderline statement suppression	Three configurable settings
Standard measurements	• Ten interval, duration, and axis measurements
Extended measurements	• Configurable QT correction method • 46 measurements of morphology analysis in each of the 18 leads • 21 parameters of rhythm analysis
Reasons	Selectable explanations of all interpretive statements
Nomenclature	Aligned with 2007 AHA/ACCF/HRS Recommendations, Part II[1]
STEMI clinical decision support	
Graphical ST presentation	• Five ECG reports with polar ST Maps • Frontal and transverse planes
Unique right heart statements	Nine statements called by right-chest leads
Unique posterior MI statements	16 statements called by posterior leads
Age and gender criteria	Based upon 2009 AHA/ACCF/HRS Recommendations, Part VI: Acute Ischemia/Infarction[2]
STEMI-CA (Culprit Artery)	• Criteria that suggest any of four probable sites of the occluded coronary artery • Based upon 2009 AHA/ACCF/HRS Recommendations, Part VI[2]
Critical values	Highlights four conditions requiring immediate clinical attention

Figure 2. *The ECG machine used in the study, its software, and specifications.*

chosen at random, using a computerized random number generator. These were grouped into sets of 20 ECG each, making 10 sets. A further 10 ECG traces were chosen at random to serve as control ECGs and to be placed in every set. Thus, we ended up with 10 sets, each made up of a total of 25 traces: 20 test ECG traces, and 5 control ECG traces.

The ECG sets were uploaded to ClassMarker®, an internet-based examination software, ready to be emailed to our test subjects made up of board-certified emergency medicine attending physicians (EMPs). Each EMP subject was presented with a link to randomly access 2 sets of ECGs, a total of 50 traces. The ClassMaker® software would then present each trace with its respective one-line of history (Figure 1) to the EMP subject, asking them to choose the best course of action for the displayed ECG trace. As noted previously, the subjects were allowed to choose one of three possible actions: 1) put the patient in the waiting room, they can wait, 2) I need to see this patient in the Triage to decide what to do with them, and 3) admit this patient to a monitored ED bed, they need immediate assessment.

The ClassMarker® software then collected the EMP subject's response for each trace displayed, a total of 50 responses for 50 traces in total. These responses would later be analyzed to see if the simple classification allocated by the ECG machine (The PageWriter® TC50 and its software, Figure 2) correlated with the EMP subject's decision, thus establishing if such a machine classification is reliable for use, in place of the need for actual physician opinion and initial assessment.

Each EMP subject enrolled submitted their responses for 2 ECG sets, a total of 40 test ECG traces and 10 control ECG traces. The inter-subject agreement and correlation coefficient kappa (k) was calculated for all subjects enrolled and in all possible permutations. The proportions of the responses made by the subjects for each ECG trace classified as "Normal," "Otherwise normal," "Borderline," and "Abnormal" were calculated.

Results

In the time frame specified, a total of 3,149 patients were registered in the EMR at KAMCJ. Of those, only 452 had ECGs done for them in the triage. Of the 452 ECGs 200 were chosen randomly as specified in the methods section, and a further 10 were chosen for use as the control. Ten ECG sets were created, each set is made of 25 ECG traces (5 of each classification type: 5 classified as normal, 5 as otherwise normal, 5 as borderline, 5 as abnormal, and 5 as control). A link was sent to a total of 25 board-certified EMPs (the EMP subjects) chosen at random that fulfil the inclusion/exclusion criteria. Each subject had access to 2 sets of ECGs (a total of 50 traces), along with instructions informing them of the objective of the study.

Only 21 subjects, out of the original 25 contacted, responded. Data from those enrolled in 21 subjects were collected and included in our analysis. The inter-subject correlation was first worked out as the kappa coefficient (k). The k-values were done for all subject permutations and ranged from 0.2394 to 0.4281. The average k was worked out to be 0.315 with an average standard deviation of 0.187. This level of k suggests a fair to moderate agreement between the enrolled EMP subjects.

Interestingly, our results show that of those ECGs, classified as "Normal" by the machine, EMPs only sent 47.6% to wait in the waiting room. While 14.8%, of the "Normal" ECGs were admitted to a critical bed to be seen immediately, by the EMP subjects (Table 1). A similar response was found for those ECGs classified as "Otherwise Normal." 49% were sent to the waiting room, while up to 20% were put on critical beds for immediate assessment.

In contrast, when looking at the ECGs classified as "Abnormal," which one would expect to end up on ED critical beds immediately, only 44.3% were placed there. While 27.1% of the "Abnormal" ECG traces were actually sent to the waiting room (Figure 3). Looking at the EMP's responses, we find that as shown in Figure 4, of those they decided to put in the waiting room, only 29% were classified as "Normal." More importantly, of those that needed to be admitted to an ED bed immediately, 34% where either "Normal" or "Otherwise normal."

A further analysis was done to see if clinical experience of the EMP subjects (represented by the number of years of practice) played any role in the response proportions or the correlation of the machine classification and the EMP decision. As shown in Figures 5 and 6 senior physicians (those that have been practicing for more than 5 years) and those that have finished their boards with in the last 5 years showed similar responses. Although there was a tendency for the senior physicians to admit more of the "Abnormally" classified ECGs to ED beds for immediate evaluation, that difference had an insignificant p-value of 0.312 and thus could be attributed to random chance.

Table 1. *EMP classification of the ECG samples.*

Machine classification	Choose to put patient in waiting room		Choose to see patient in triage		Choose to put patient on an ED bed immediately		Total
	n	%	*n*	%	*n*	%	
Normal	100	47.6%	79	37.6%	31	14.8%	210
Otherwise-normal	103	49.0%	65	31.0%	42	20.0%	210
Borderline	81	38.6%	83	39.5%	46	21.9%	210
Abnormal	57	27.1%	60	28.6%	93	44.3%	210

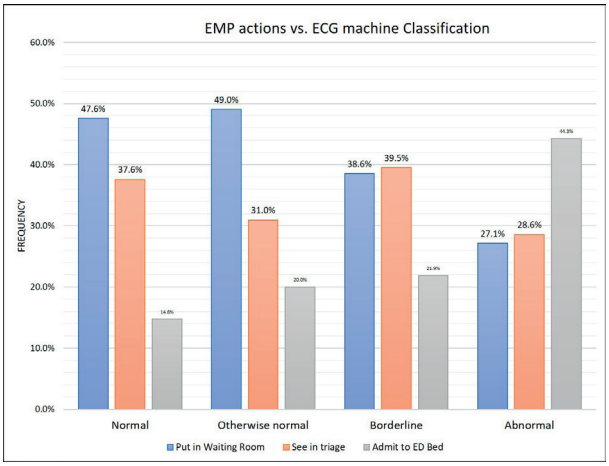


Figure 3. *EMP versus ECG machine results.*

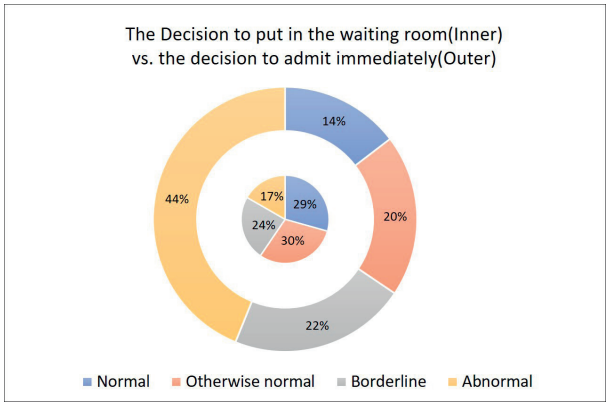


Figure 4. *Decision to put in the waiting room (inner) versus the decision to admit immediately (outer).*

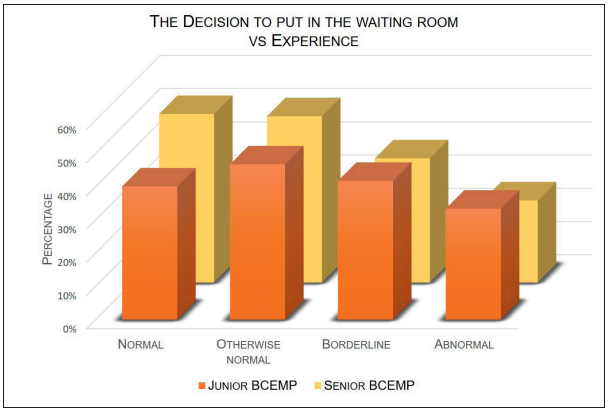


Figure 5. *Decision to put in the waiting room versus experience.*

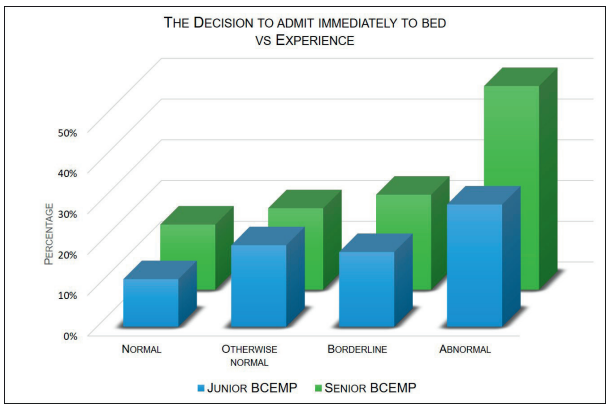


Figure 6. *Decision to admit immediately to bed versus experience.*

Discussion

This study shows similar results as the previous evidence [5,7]. To date, computer ECG interpretation has never been trusted. This is simply because there is so much that goes into the medical assessment of a patient presenting to the ED with chest pain. The process may seem simple on the surface, however, so many unseen pieces of information factor into the decision of a physician to decide to put the patient immediately in an ED bed or to let them wait in the waiting area. As computer algorithms improve with leaping technological advancement, we are sure that there will come a time when the software may eventually develop to encompass all these bits of data, including the look on the patient's face, a muscle twitch, or a facial grimace. However, that time is still in the future. As far as we can tell, the simple classification of the patient's ECG trace as "Normal," "Otherwise normal," "Borderline," or "Abnormal" does not mirror the expert judgment of a board-certified EMP in any close measure.

Should this simple classification have worked as a simple screening test where an ECG labeled by the machine as "Normal" could be reliably moved to the waiting area, this would have saved the emergency physician a lot of interruption time. Policies could have been changed to use this by triage nurses, instructing them to only ask for the physician opinion on the flow of the patient if the ECG was labeled "Abnormal," and that all "Normal" labeled traces would find their way to the waiting room. Alas, this was not the case.

It is very evident that expert opinion is missed by the simplest of machine classification, mainly being normal or not. The lack of correlation is not only of note, but

in some cases, it almost reaches 50% difference in both normal and abnormal ECG traces. Furthermore, it would seem that this mismatch between machine classification and professional opinion does not seem to change based on years of experience. It is therefore painfully clear that for the time being, nurses and other healthcare personnel manning the triage, will need to keep bothering attending physician, who already has a thinly stretched attention span, during their busy shift to make a decision on the patient currently lying in the triage room connected to the ECG machine complaining of chest pain, until a more reliable artificial intelligence can be programmed into the ECG machines of the future to allow them to sensitively and reliably take a decision in the physician's stead.

Despite its strengths, our study suffers from multiple limitations as well. The ECG machine and software is only one of the many models available, should there be a better model that offers better software to mirror the professional decision of the EMP this was not accounted for or studied. The model used in this study was the most common model used in the MOH hospitals in Saudi Arabia, but not the only model used. The actual final diagnosis and disposition of the patients included in this study were not included in the analysis, which will serve to validate the EMP's response if they were correct or not. Although patient vitals do play a very significant part in the decision process, these were not included in the one-line history given with each ECG to the study subjects. However, this was done to mimic what is actually occurring in real life, and the nurse does not always include the vital signs when asking for the attending's opinion on an ECG trace. Rather, it is the physician that usually asks for the vitals and this was supposed to be the decision where the attending would opt to see the patient in the triage to decide on what to do.

List of Abbreviations

ECG	Electrocardiogram
ED	Emergency department
EMR	Electronic medical record
KAMCJ	King Abdullah Medical Complex Jeddah
MOH	Ministry of Health

Conflict of interests

The authors declare that there is no conflict of interest regarding the publication of this article.

Funding

None.

Consent to participate

Written informed consent was obtained from all the participants.

Ethical approval

Ethical approval was granted by IRB Research Committee, approval no. A01319, given on March 12, 2022.

Author details

Afnan S. Sharourou[1], Omar T. Idriss[1], Ahmed L. AlManzalawi[2], Amro M. Tayeb[3], Thamer A. AlGhamdi[4], Anas Fouad Hamam[5,6]

1. Saudi Arabian Board of Emergency Medicine Joint Residency Training Program, Jeddah, Saudi Arabia
2. Department of Emergency Medicine, King Abdullah Medical Complex, MOH, Jeddah, Saudi Arabia
3. Saudi Arabian Board of Internal Medicine Residency Training Program, Jeddah, Saudi Arabia
4. Consultant of Emergency Medicine Residency Director, King Abdullah Medical Complex, MOH, Jeddah, Saudi Arabia
5. Consultant of Emergency Medicine, King Fahad Armed Forces Hospital, Jeddah, Saudi Arabia
6. Emergency Medicine Joint Residency Training Program Director, Jeddah, Saudi Arabia

References

1. Gulati M, Levy PD, Mukherjee D, Amsterdam E, Bhatt DL, Birtcher KK, et al. 2021 AHA/ACC/ASE/CHEST/SAEM/SCCT/SCMR guideline for the evaluation and diagnosis of chest pain: a report of the American College of Cardiology/American Heart Association Joint Committee on clinical practice guidelines. Circulation. 2021;144(22):e368–454. https://doi.org/10.1161/cir.0000000000001029
2. Chisholm CD, Collison EK, Nelson DR, Cordell WH. Emergency department workplace interruptions: are emergency physicians "interrupt-driven" and "multitasking"? Acad Emerg Med. 2000;7(11):1239–43. https://doi.org/10.1111/j.1553-2712.2000.tb00469.x
3. Allard J, Wyatt J, Bleakley A, Graham B. "Do you really need to ask me that now?": a self-audit of interruptions to the 'shop floor' practice of a UK consultant emergency physician. Emerg Med J. 2012;29(11):872–6. https://doi.org/10.1136/emermed-2011-200218
4. Blocker RC, Heaton HA, Forsyth KL, Hawthorne HJ, El-Sherif N, Bellolio MF, et al. Physician, interrupted: workflow interruptions and patient care in the emergency department. J Emerg Med. 2017;53(6):798–804. https://doi.org/10.1016/j.jemermed.2017.08.067
5. Stopa M, Sevenathan H, Bogusławski M, Pałasz I. Assessment of ECG interpretation skills among Polish medical students, nursing, emergency medicine and English division medical students: PS044. Porto Biomed J. 2017;2(5):231–2. https://doi.org/10.1016/j.pbj.2017.07.129
6. Werner K, Kander K, Axelsson C. Electrocardiogram interpretation skills among ambulance nurses. Eur J Cardiovasc Nurs. 2016;15(4):262–8. https://doi.org/10.1177/1474515114566158
7. Hamam AH, AlNofaiey YH, AlAlayani AM. Adequacy of electrocardiogram reading and interpretation among emergency medicine residents. Am Res J Emerg Intensive Crit Care. 2016;1(1):1–7.