Exploring decision making 'noise' when interpreting the electrocardiogram in the context of cardiac cath lab activation

Peace, A., Al-Zaiti, S., Finlay, D., McGilligan, V. E., & Bond, RR. (2022). Exploring decision making 'noise' when interpreting the electrocardiogram in the context of cardiac cath lab activation. *Journal of Electrocardiology*, *73*, 157-161. https://doi.org/10.1016/j.jelectrocard.2022.07.002

Published in:
Journal of Electrocardiology

Publication Status:
Published (in print/issue): 31/08/2022

DOI:
https://doi.org/10.1016/j.jelectrocard.2022.07.002

Document Version
Author Accepted version

Exploring decision making 'noise' when interpreting the electrocardiogram in the context of cardiac cath lab activation

Aaron Peace[1], Salah S. Al-Zaiti[2], Dewar Finlay[3], Victoria McGilligan[3], Raymond Bond[3]
[1]Clinical Translational Research and Innovation Centre, Northern Ireland, UK
[2]University of Pittsburgh, Pittsburgh PA, USA
[3]Ulster University, Northern Ireland, UK

Abstract
In this commentary paper, we discuss the use of the electrocardiogram to help clinicians make diagnostic and patient referral decisions in acute care settings. The paper discusses the factors that are likely to contribute to the variability and noise in the clinical decision making process for catheterization lab activation. These factors include the variable competence in reading ECGs, the intra/inter rater reliability, the lack of standard ECG training, the various ECG machine and filter settings, cognitive biases (such as automation bias which is the tendency to agree with the computer-aided diagnosis or AI diagnosis), the order of the information being received, tiredness or decision fatigue as well as ECG artefacts such as the signal noise or lead misplacement. We also discuss potential research questions and tools that could be used to mitigate this 'noise' and improve the quality of ECG based decision making.

Introduction
Primary percutaneous coronary intervention (PPCI) is the act of emergent recanalisation of an acute thrombotic occlusion of a coronary artery which is often first diagnosed using the 12-lead electrocardiogram (ECG). And whilst conventional ST elevation myocardial infarction (STEMI) criteria are typically used to detect a suspected coronary occlusion, there is an emerging approach for detecting coronary occlusions that is referred to as occlusion myocardial infarction (OMI) (Aslanger et al., 2021). The patient's journey invariably begins upon the intervention of a paramedic who obtains an ECG from the patient to ascertain whether their symptoms are due to a coronary occlusion. This electrical biomarker has been used in clinical practice for over 70 years and is the critical determinant of the subsequent journey the patient takes. The STEMI criteria typically dictates whether the patient is taken immediately to a PPCI enabled centre bypassing the emergency department or whether they are directed to their closest emergency room. Whilst this is a well established process in most modern healthcare systems, evidence shows that there continues to be significant inefficiencies particularly with respect to, what are described as, 'false activations' (Tolles et

al., 2020). In this commentary paper, we will discuss the factors that influence the clinical decisions which remain less than ideal.

Decision making

In general, the ability of humans to make decisions varies widely and is influenced by many factors. The ECG is an established, inexpensive and noninvasive tool to help clinicians make decisions. A key example is the use of the ECG to inform whether the patient requires emergency PPCI. The ECG itself is a recording of the electrical phenomena that is exhibited by the heart as acquired from the body surface. This alone tells us that the ECG may lack some detail given that it is observing cardiac electrical activity from a 'distance'. Moreover, whilst the ECG presents electrical activity, this electrical activity is interpreted to infer the health of the mechanics, structure and functioning of the heart, including any changes to the ECG. In this context, the ECG is a crucial tool for deciding whether someone is having an acute myocardial infarction. However, we need to be cognisant of the factors that could affect the quality of the decision making when reading the ECG. These factors can be described as adding 'noise' to the decision making process. For the engineer, the term 'noise' typically refers exclusively to issues with electrical and related interference that affects the morphology of the ECG signal itself. In this article, we broaden the definition of 'noise' to be factors that can provide contamination to the decision making process. However, we do make some reference to the effects of noise in the signal itself in the later sections.

Noise factors

One of the greatest threats to providing high quality healthcare is poor clinical decision making. There are multiple factors influencing human decision making, many occurring simultaneously such that 'noise' is produced which is undesirable. Kahneman, Sibony and Sunstein explain, *"wherever there is judgement, there is noise, and more of it than you think"* (Kahneman et al., 2021).

Varied interpretation accuracy and disagreement

A human factor that contributes to noise in ECG interpretation is the varied competence of different decision makers (ECG interpretation accuracy) as well as intra and inter rater reliability. Intra rater reliability is the agreement rate with yourself after having interpreted the same ECGs or patient cases at different time points. Inter rater reliability is the agreement rate between different interpreters of the same ECGs (or patient cases). These metrics can be computed using agreement rates (in percentage) or kappa statistics. A recent meta-analysis carried out by Cook et al. (2020) demonstrates that there is room for improving the accuracy of interpreting ECGs for both students and doctors. It has also been pointed out that between 4% and 33% of ECG interpretations have significant errors (Salerno et al., 2003). However, Lim et al. (2015) reported that physicians have good agreement regarding the identification of the J point and the measurement of ST amplitudes. Despite this, however, McCabe et al. (2013) found that there was poor agreement (kappa=0.33) between physicians when interpreting ECGs to detect STEMI. This illustrates the amount of noise in clinical decision making with the ECG. And given the lack of standardisation in ECG training, we are potentially sustaining the noise instead of mitigating it.

The algorithms

Of course, there are a myriad of other noise factors. One being the influence of the machine diagnosis on the human interpreter. This is important given that a recent study by Faramand et al. (2021) shows that machine diagnosis only has a sensitivity of 63% for detecting STEMI in patients with chest pain. If the human reader initially considers the machine's interpretation before undertaking their own unbiased interpretation, then they could be prone to 'automation bias' which is where they could complacently accept the automated suggestion (Bond et al., 2018). This is perhaps similar to anchoring, confirmation bias and acquiescence bias which are well known cognitive biases in research. It is perhaps easier to agree than to disagree with an 'external suggestion'. With this in mind, it might be important to discover and engineer the optimal order of information (or information flow) for physicians to receive when making these decisions (see Figure 1). For example, perhaps reading the ECG without knowing the machine diagnosis is a better approach, and only after this, would the automated interpretation be revealed. Moreover, it is also possible that the automated ECG measurements might prevent interpreters from precisely measuring or double checking amplitudes and interval measurements themselves. This does present a research question, namely, does the order of the information being received affect how physicians make their final decisions?

The flow of information

We know from the 'serial position effect' (Wong, 2021), that the order of items in a list does influence what we store in our working memory, i.e. the primacy and recency effects tell us that we are more likely to remember the first and last items in a list. Cairns et al. (2017) developed a decision support system that goes further and controls the order of which ECG signals and leads the reader interprets. In this way, the system dictates the order of the information in the ECG interpretation process over multiple interactive screens. This ensures that the reader is systematic and considers all signals and diagnoses. When using this system, the reader only sees the full 12-lead ECG and the machine diagnosis at the end of the ECG interpretation process as opposed to the start of the process. This approach was called an 'interactive progressive based interpretation' and could have the potential to reduce the noise in ECG based decision making. In the field of forensic science, Dror et al. (2021) refers to a similar approach called 'linear sequential unmasking' (LSU). And, in their paper (Dror et al., 2021), they suggest that managing the order of the information flow can reduce bias. Hence, this approach could help mitigate the problem of 'making your first impression - the only lasting impression'. Kahneman, Sibony and Sunstein (2021) refer to what they call, a 'decision hygiene' strategy, whereby it is best to initially focus on the facts and deliberate reasoning, and postponing your intuition (i.e. delaying your 'gut reaction' when first seeing the ECG). Perhaps this means that we should be deliberately systematic at the start of the ECG interpretation process and after that, only then should we engage our automatic pattern recognition skills to ECG interpretation.

Time pressures, stress and tiredness

In addition to poor accuracy and agreement, there are potentially other factors that add additional noise to the decision making process. This includes reading ECGs in time critical scenarios when 'time is muscle'. ECG misinterpretation during prehospital transport has been shown to be one of the strongest predictors of undertriage in patients with potential acute myocardial infarction (Faramand et al., 2019a). There are many research questions that need to be further explored. For example, does time pressure and stress influence ECG interpretation and the referral decision? Are physicians more likely to accept the machine's ECG interpretation when they are under time pressure? Perhaps stress and tunnel vision could affect the capacity to recognise and process all relevant information. Nevertheless, there are many other research questions that could be proposed, for example, does the amount of time that has lapsed on a physician's working hospital shift affect their ability to accurately read ECGs and make referral decisions? Are physicians less accurate in reading ECGs at the end of a long shift when compared to the start of their shift? A recent study shows that the peak in STEMI encounters by paramedics falls within the daily surge in chest pain transports (Faramand et al., 2019b).

Decision fatigue

We should also consider 'decision fatigue', which is when the ability of making decisions is impaired due to the number of decisions that have already been made. Hence, a relevant question is: how does such 'decision fatigue' affect the quality and accuracy of ECG interpretations? Allan et al. (2019) showed that decision fatigue can result in nurses being more "conservative" in decision making. This prompts us to think whether decision fatigue could result in healthcare professionals being more prone towards simply accepting the machine diagnosis (automation bias).

Type of error being avoided

In addition, perhaps the culture, personalities and the general environment can also affect the quality of the decision making. For example, is there more fear in making type 1 or type 2 errors. To explain, perhaps a member of staff predominantly fears missing a heart attack patient whereas another member of staff might fear sending too many false referrals - or indeed is the balance of 'fear' between making these two types of errors more dynamic. A multidisciplinary team is involved in the decision making that takes place during the PPCI pathway, including the nurse activator who plays a vital role in the activation of the pathway. Interdisciplinary teamwork and collaboration have been previously discussed as key factors involved in the decision making ability of nurse activators (Clayton, 2019). Personalities of each team member may play a role in the decision making and relationships between these personalities could perhaps influence decisions. For example, what is the dynamic and conformity to the highest paid person's opinion (HiPPO)?

Technical factors

There are of course other technical factors that contribute to the variability of ECG based decision making. These include the level of signal noise in the ECG (e.g. mains noise [50/60Hz], baseline wander etc.), the various settings of different ECG machines, the different algorithmic logic/computer programs from different manufacturers, different filter settings and of course electrode misplacement which could also alter the signal (Rjoob et al., 2020). There may also be problems and challenges when electronically transmitting ECGs (Al-Zaiti et al., 2013), for example with incomplete data in the context of telemedicine. All of these noise factors (as illustrated in Figure 2) could have a significant impact on the quality of decisions, especially if they are compounded. We should consider the compound effect of the accumulation of marginal errors as the inverse of the 'aggregation of marginal gains' (Clear, 2018).

Based on this brief commentary, we have enumerated potential noise mitigation approaches in Table 1 which could be used to reduce the noise and improve ECG based decision making. Moreover, Table 2 presents potential prospective research studies that could be carried out to improve our understanding of ECG based decision making and how we can improve this process.

Conclusion

There are a plethora of factors that can contribute to noise when making decisions using the ECG. More research is required to understand the level of noise that is present in departments that use the ECG to make decisions. Kahneman, Sibony, and Sunstein (2021) present a framework for undertaking a 'noise audit'. This focuses on developing an experiment to understand the variability in human judgements when collecting decisions from different decision-makers using well designed representative cases (for example patient vignettes). It also involves determining what levels of noise is expected/anticipated by the decision makers and what levels of noise is considered acceptable, as well as defining the cost of misjudgements (or misinterpretations in the case of the ECG). However, as discussed, there are many other factors such as the different filter settings and lead misplacement which can also add additional variability to the decision making process. Hence, it might be a good idea to develop a standard framework to assess the decision making variability and 'noise' in a department that uses the ECG to make decisions. In conclusion, the more we understand the problem and the human factors in ECG based decision making, the more likely we are to provide a solution that can improve the quality and consistency of clinical decision making.

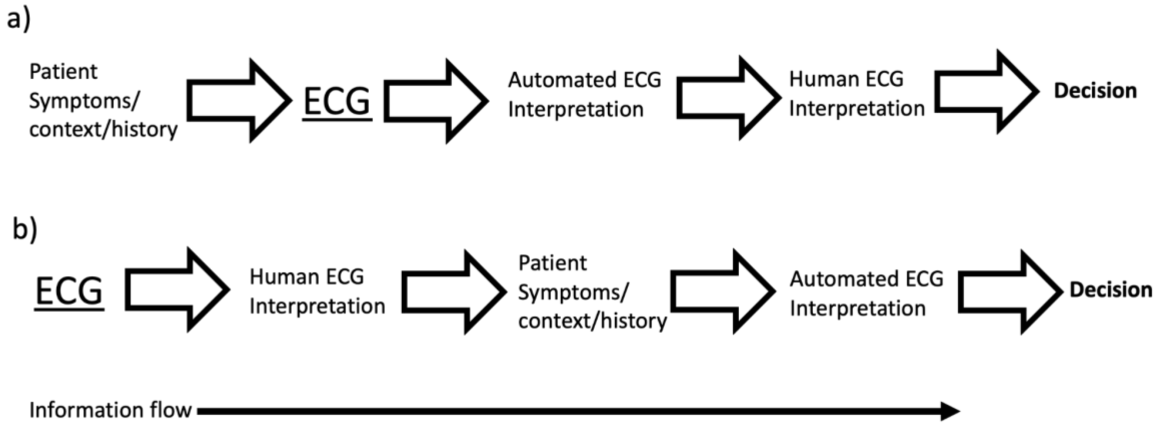

Figure 1. Illustration of an example research question (does the order of the information flow affect how staff read the ECG?). a) perhaps a typical flow of information and context for when a healthcare professional interprets the ECG, b) an alternative paradigm for the flow of information that may mitigate biases in the ECG interpretation.

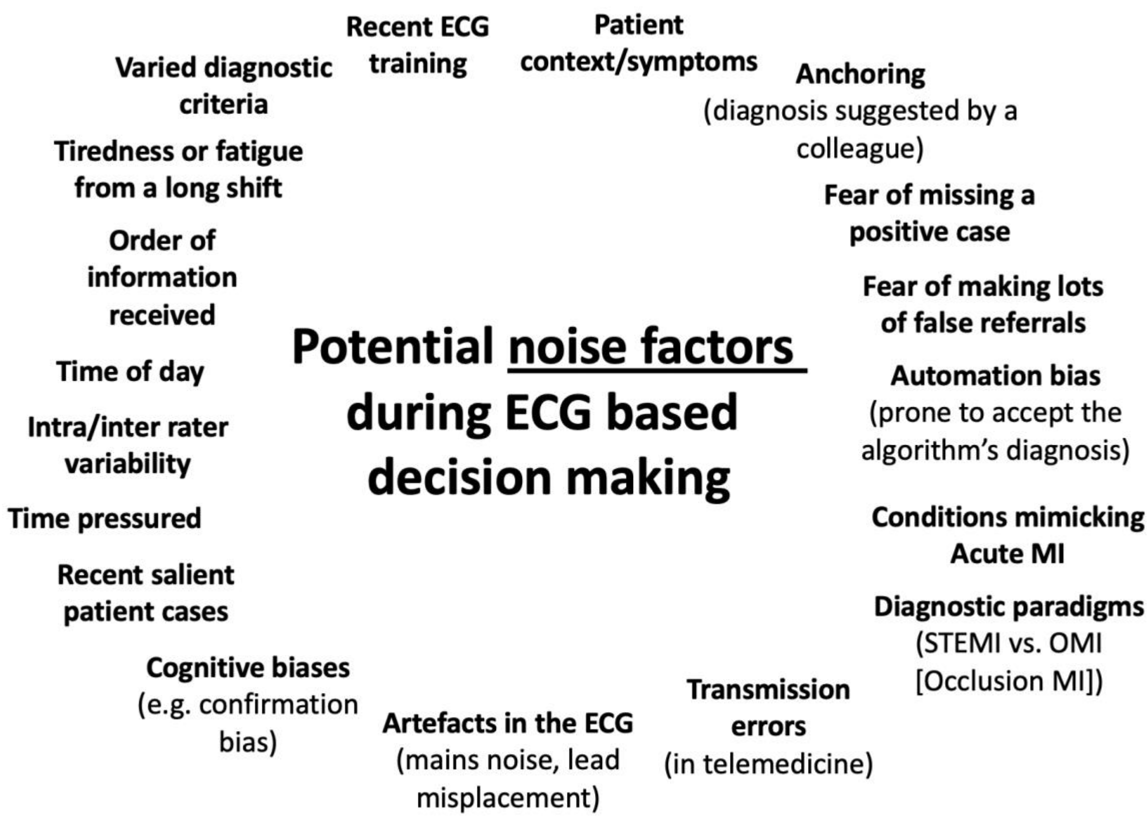

Figure 2. Potential noise factors during ECG based decision making.

Table 1. Example of noise mitigation approaches to improve the decision making with ECG based decision making.

Noise mitigation	Rationale
Interpret the ECG using a checklist or using a system where possible.	This ensures that each ECG is systematically interpreted.
Delay using system 1 thinking* and avoid relying on immediate intuition and use reasoning before intuition. ** System 1 thinking is when we use our automatic intuition/first impression to make a decision - i.e. our 'knee jerk' reaction* (Kahneman et al., 2011)	This is related to the former mitigation. This technique has been called a "Decision Hygiene" strategy (Kahneman et al.., 2021)
Do not read the automated diagnosis until you have interpreted the ECG yourself.	This avoids automation bias and anchoring and allows the physician to read the ECG without having being biassed by the influence of a suggested automated interpretation.
Design and engage in a standardised high quality frequent ECG training programme.	This ensures that each member of staff has the same training and if the training is regular then this could combat the Ebbinghaus 'forgetting curve' (Murre and Dros, 2015).
Measure the inter and intra rater reliability of staff and inform each member of their performance and their intra rater reliability.	Allows staff to know what noise exists in their centre and what their own intra rater reliability is. Being aware of this could help staff engage with training.
Regular demonstrations of proper electrode placement and case studies of electrode misplacement.	Ensures ECGs are properly recorded by different members of staff and case studies demonstrate the potential impact on decisions.
Undertake decision making research to detect noise factors in ECG based decision making.	For example, studying the accuracy of ECG interpretation at the start and end of shifts or a similar study could aid noise mitigation programmes and help with service quality improvement
Ask for help when needed but request an independent assessment without anchoring colleagues based on your suggested	This provides independent assessments and seeks a reliable consensus when it is needed.

interpretation.	
Develop new algorithms that accurately detect coronary occlusions (Faramand et al., 2021)	A more accurate and approved algorithm, perhaps based on modern artificial intelligence techniques could be used. For example, the use of deep learning to detect coronary occlusions from ECGs where the algorithm is trained using labels that have a greater ground truth i.e. based on immediate angiographic findings.

Table 2. Example potential research questions that could add to the body of knowledge on the quality of ECG based decision making.

Research question	**Impact**
Does ECG interpretation accuracy change depending on how tired the physician is?	MIght inform or optimise the frequency of breaks or times when decisions should be checked.
Are there fewer false referrals to PPCI after ECG interpretation training?	Would inform the extent of the need for standardised ECG training and its frequency.
Does delaying exposure to the automated diagnosis and automated ECG measurements improve decision making?	Informs the design of how the ECG and automate analysis should be presented and how the optimal order of the information.
With knowledge of one's own intra rater reliability and accuracy influence their decision making?	Would inform whether we should all engage in a personal audit to provide self-knowledge and insight into our own interpretation variability.
How can lead misplacement affect automated diagnoses and human ECG interpretation?	Would inform the need for algorithms and better training to detect these errors.
Does knowing the patient history and symptoms improve or negatively bias ECG interpretation?	Informs whether the ECG should initially be interpreted with minimal information about the patient.
Does the order of the information flow affect how staff read the ECG?	Would inform whether there should be an optimal sequence of information to process when making a decision in PPCI.
To what extent does a time pressured environment and/or hospital distractions	Would inform the expected error in decisions under time pressure and whether

impact ECG interpretation performance?	the extent of the error is acceptable.
What levels of inter/intra rater reliability in ECG interpretation do staff believe is acceptable/tolerable? What is the difference between actual versus perceived magnitude of inter/intra rater reliability in ECG interpretation?	Would provide research on the awareness or lack of awareness of the quality of decision making and help calibrate efforts to reduce any unexpected error.
Can deep learning reduce the 'noise' in ECG based decision making and provide more consistent and accurate interpretations? What would be the consequences of using deep learning - would there be a reduction in the physician's competence in reading ECGs?	Would inform whether deep learning would be used in clinical practice and whether we need to mitigate the risk of attenuating the physician's competence in reading ECGs.

Acknowledgements

Aaron Peace, Raymond Bond and Victoria McGilligan are supported by the European Union's INTERREG VA programme (managed by Special EU Programmes Body - SEUPB) under the project: 'Centre for Personalised Medicine—Clinical Decision Making and Patient Safety'. The views and opinions in this study do not necessarily reflect those of the European Commission or the Special EU Programmes Body (SEUPB).

References

Allan, J.L., Johnston, D.W., Powell, D.J., Farquharson, B., Jones, M.C., Leckie, G. and Johnston, M., 2019. Clinical decisions and time since rest break: An analysis of decision fatigue in nurses. Health Psychology, 38(4), p.318.

Al-Zaiti, S.S., Shusterman, V. and Carey, M.G., 2013. Novel technical solutions for wireless ECG transmission & analysis in the age of the internet cloud. Journal of electrocardiology, 46(6), pp.540-545.

Aslanger, E.K., Meyers, P.H. and Smith, S.W., 2021. STEMI: A transitional fossil in MI classification?. Journal of Electrocardiology, 65, pp.163-169.

Bond, R.R., Novotny, T., Andrsova, I., Koc, L., Sisakova, M., Finlay, D., Guldenring, D., McLaughlin, J., Peace, A., McGilligan, V. and Leslie, S.J., 2018. Automation bias in medicine:

The influence of automated diagnoses on interpreter accuracy and uncertainty when reading electrocardiograms. Journal of electrocardiology, 51(6), pp.S6-S11.

Cairns, A.W., Bond, R.R., Finlay, D.D., Guldenring, D., Badilini, F., Libretti, G., Peace, A.J. and Leslie, S.J., 2017. A decision support system and rule-based algorithm to augment the human interpretation of the 12-lead electrocardiogram. Journal of electrocardiology, 50(6), pp.781-786.

Clayton, W.R., 2019. Overcoming Barriers impeding nurse activation of rapid response teams. OJIN: The Online Journal of Issues in Nursing, 24(3).

Clear, J., 2018. Atomic habits: An easy & proven way to build good habits & break bad ones. Penguin.

Cook, D.A., Oh, S.Y. and Pusic, M.V., 2020. Accuracy of physicians' electrocardiogram interpretations: a systematic review and meta-analysis. JAMA internal medicine, 180(11), pp.1461-1471.

Dror, I.E. and Kukucka, J., 2021. Linear sequential unmasking–expanded (LSU-E): a general approach for improving decision making as well as minimizing noise and bias. Forensic Science International: Synergy, 3.

Faramand, Z., Frisch, S.O., DeSantis, A., Alrawashdeh, M., Martin-Gill, C., Callaway, C. and Al-Zaiti, S., 2019a. Lack of significant coronary history and ECG misinterpretation are the strongest predictors of undertriage in prehospital chest pain. Journal of Emergency Nursing, 45(2), pp.161-168.

Faramand, Z., Frisch, S.O., Martin-Gill, C., Landis, P., Alrawashdeh, M., Al-Robaidi, K.A., Callaway, C.W. and Al-Zaiti, S.S., 2019b. Diurnal, weekly and seasonal variations of chest pain in patients transported by emergency medical services. Emergency Medicine Journal, 36(10), pp.601-607.

Faramand, Z., Helman, S., Ahmad, A., Martin-Gill, C., Callaway, C., Saba, S., Gregg, R.E., Wang, J. and Al-Zaiti, S., 2021. Performance and limitations of automated ECG interpretation statements in patients with suspected acute coronary syndrome. Journal of Electrocardiology, 69, pp.45-50.

Kahneman, D., 2011. Thinking, fast and slow. Macmillan.

Kahneman, D., Sibony, O. and Sunstein, C.R., 2021. Noise: A flaw in human judgment. Little, Brown.

Lim, H.C.S., Salandanan, E.A., Phillips, R., Tan, J.G. and Hezan, M.A., 2015. Inter-rater reliability of J-point location and measurement of the magnitude of ST segment elevation at the J-point on ECGs of STEMI patients by emergency department doctors. Emergency Medicine Journal, 32(10), pp.809-812.

McCabe, J.M., Armstrong, E.J., Ku, I., Kulkarni, A., Hoffmayer, K.S., Bhave, P.D., Waldo, S.W., Hsue, P., Stein, J.C., Marcus, G.M. and Kinlay, S., 2013. Physician accuracy in interpreting potential ST-segment elevation myocardial infarction electrocardiograms. Journal of the American Heart Association, 2(5), p.e000268.

Murre, J.M. and Dros, J., 2015. Replication and analysis of Ebbinghaus' forgetting curve. PloS one, 10(7), p.e0120644.

Rjoob, K., Bond, R., Finlay, D., McGilligan, V., Leslie, S.J., Rababah, A., Guldenring, D., Iftikhar, A., Knoery, C., McShane, A. and Peace, A., 2020. Machine learning techniques for detecting electrode misplacement and interchanges when recording ECGs: A systematic review and meta-analysis. Journal of Electrocardiology, 62, pp.116-123.

Salerno, S.M., Alguire, P.C. and Waxman, H.S., 2003. Competency in interpretation of 12-lead electrocardiograms: a summary and appraisal of published evidence. Annals of Internal Medicine, 138(9), pp.751-760.

Tolles, J., Bosson, N., Kaji, A.H., Henry, T.D., French, W.J., Gausche-Hill, M., Andruss, K., McNeil, N., Nakkim, E.C., Thomas, G.S. and Gunderson, M.R., 2020. The effect of implementation of the American Heart Association Mission Lifeline preact algorithm for prehospital cardiac catheterization laboratory activation on the rate of "false positive" activations. Prehospital and disaster medicine, 35(4), pp.388-396.

Wong, E., 2021. Serial Position Effect: How To Create Better User Interfaces. Interaction Design Foundation, https://www.interaction-design.org/literature/article/serial-position-effect-how-to-create-better-user-interfaces