Automated Screening of Patients for Dietician Referral*

Kamran Soomro and Elias Pimenidis

Computer Science Research Centre, University of the West of England, Bristol, UK
{kamran.soomro, elias.pimenidis}@uwe.ac.uk

Abstract. Critical Care Units (CCU) in a hospital treat the severely sick patients that need constant monitoring and close medical attention. Feeding patients, enteral feeding in particular, is a critical and continuous process. Monitoring patients, managing their feeding and referring to a dietician is a key factor in CCUs. Screening patients for referral to a dietician in a CCU is an error-prone and complicated task. One of the main challenges in this regard is that the data needed to screen patients is scattered among many different variables and textual forms. The number of patients being treated in the CCU is also a significant problem since it becomes difficult for the staff to keep track of the needs of all patients. Therefore, an automated screening tool can support effectively the feeding process and contribute considerably towards improving the quality and consistency of patient care. In this paper we present early stages of a project that aims at using machine learning techniques to help CCU consultants to automatically screen patients for dietician referral.

Keywords: automated patient screening · CCU screening · Dietician referral · Patient Data Processing

1 Introduction

This paper presents the early stages of a project that aims at using Artificial Intelligence to analyse data collected in the critical care units (or intensive care units) of a National Health Service (NHS) hospital in the UK. Critical Care Units (CCUs) are where the sickest patients are admitted and where large amounts of detailed clinical data are collected (usually hourly) for the duration of the patient's stay. Most CCUs, have moved from paper to clinical information systems to capture data from the patients' monitor, ventilator and other equipment into an extensive database. Currently, CCUs are not using these systems and data to their advantage, with much of the captured data left unanalysed. In an era of limited NHS resources, these digital resources have the potential to both optimise patient outcomes and make systems more efficient and effective.

Data analysis and machine learning systems have been used widely in CCUs, primarily to automate processes where problems are well known and fully defined, and there the deliverables to efficiency and accuracy achieved are well within the expected levels [1, 7].

* Supported by the University of the West of England, Bristol.

CCUs aim to optimise the patients' survival, clinical outcomes and to reduce harm caused by therapies. All CCUs have recognised targets to improve outcomes, such as maintaining an optimal sedation level, maintaining lung volumes delivered by the ventilator within a specific range and delivering a minimum amount of nutrition to patients whilst they are critically ill. Yet, these targets are often not achieved. This study aims to use these common targets to assess the feasibility of using this clinical data routinely collected can be used to improve the achievement of these targets.

The aim of this project is to determine how routine clinical data collected in CCUs can be used to optimise patient outcomes by augmenting clinician decision making and altering clinician behaviour to optimise patient clinical outcomes.

This study will focus on optimising three common clinical targets: sedation level, mechanical ventilation and enteral nutrition in patients of all ages requiring intensive care. These interventions have been selected because they can be defined regardless of age; and are known to be poorly met in practice [9, 8]. Data produced by the clinical information system at the CCUs has been studied to determine its quality and suitability for exposure to advanced analysis techniques. Data has been filtered and cleaned to develop a suitable experimental repository that is compatible with further analysis techniques. The data will be mined to reveal potential associations that will enhance the detail , quality and significance of the information that could be passed on to the clinicians at the CCUs.

1.1 Aims and Project Plan

The collected data has been formatted in a way that allows for it to be uniform, easily stored and accessed in an efficient way. The formatting and cleaning of the data will allow for the results of processing to return outputs that are consistent in the nature and format of the output, can be interpreted consistently over the same organisational criteria, guaranteeing that outcomes are compatible throughout time. This phase will determine which of the data elements produced by the clinical information system will be useful for further processing and suitable for delivering coherent outcomes. All patient data has been anonymised in accordance to the ethics of the NHS in the UK and in accordance to the UK Data Protection Act which encompasses the GDPR.

Data will be stored in repositories that are only accessible to the members of the team of this project. The processing at this initial stage will involve data mining techniques that will aim to reveal hidden associations within the data. These will be explored as to being consistent over time, i.e. the same types of associations confirmed over different data batches. The analysis will also explore the logic of the associations revealed as to whether these are relevant to the aims of the clinicians, i.e. whether the extracted context in the data associations is suitable to support decision making in the target areas that have been identified for this project. These conclusions will allow, the members of the project team working on the analysis of the data, to carefully select among the data produced those elements that will support meaningful further processing. Thus, the volume of data that will be utilised for the next stage of the research will be those reduced to those data that will contribute directly to meeting the aims of this project, by processing via advanced Artificial Intelligence techniques.

For the next phase of the work, different machine learning techniques and algorithms will be utilised. The aim here will be to process the data items selected in the previous phase of the analysis to allow the work team and the participating clinicians to assess the efficiency and effectiveness of these techniques and algorithms. The data that has been cleaned, prepared and formatted at the initial phase will be split into training and test sections, with the formal used to train the machine learning algorithms and the latter to attempt to reproduce known results. The accuracy and efficiency, of the algorithms that will be exposed to this experimental data, will be assessed by the project team. Decisions as to the usefulness of each of the techniques explored will be made; if deemed necessary, the clinicians will provide additional patient case data to assist resolving the suitability of these techniques. On completion of this phase of processing and testing, results and conclusions will be reviewed by the project team to assess their usefulness and the potential of expanding this research further to a full scale system. Such a system will be engaged to process data in real time and aim to support decisions made by the clinicians, by drawing away the complexity of processing rich and voluminous data and providing the benefits of data associations that will support faster and more accurate decision making the CCU clinicians.

One of the most urgent needs in any CCU is that of screening and managing a patient's enteral feeding [4]. There could be complications in a patient's health that could lead to a reaction to feeding at any time during the day. These would in turn lead to decisions in changing the patient's feeding patterns, stopping feeding temporarily, or referring the patient to a dietician for further assessment. Screening is complicated due to the sources and types of data that have to be considered at very frequent intervals, the number of nursing staff available to do so and the impact it could have on the overall treatment of the patient. This has been identified as a first priority in this project.

2 The Problem

After consultation with CCU consultants, one of the problems identified to be targeted within this project is the screening of patients for referral to dieticians. The main problem in this regard is that there are numerous factors to consider. These include:

1) Patient Body Mass Index less than 18.5 kg/m^2.
2) Patient received little or no nutrition for 5 days or more.
3) Patient has been admitted in critical care $\geq$3 days and receiving enteral tube feeding as per protocol.
4) Patient received Jejunal tube feeding.
5) Patient received renal replacement therapy (intermittent or continuous) 3 days or more.

6) Patient has liver disease.
7) Patient has Pancreatitis.
8) Patient has Chyle leak.
9) Patient has significant short bowel resections and/or high output ileostomy.
10) Pressure has ulcer category 2 or above.
11) Health professionals have concerns regarding nutrition.

Determining whether a patient needs referral to a dietician requires subjective assessment by the clinicians based on these factors. The issue is further complicated by the fact that this information is not directly recorded. It either needs to be inferred from various attributes recorded in the patient history or from the text-based patient history. For example, whether a patient has liver disease or not is not directly recorded anywhere. A consultant can look at an attribute such as the patient's Albumin or Bilirubin levels and determine that the patient has liver problems, though this would still not indicate if the problems are chronic or acute [3]. That aspect is only recorded in the patient history.

The aforementioned complications present significant challenges in the automation of the nutritional screening process. Therefore, the first challenge in developing such a system is to synthesise the variables that can be fed into an appropriate learning algorithm. Different approaches could be adopted for this purpose. For example, for determining the presence of liver disease, one approach could be to perform a fuzzy word match to search the patient history for text related to liver disease. This text could then be analysing using natural language processing to determine if the patient has acute or chronic liver disease. Another approach could be to analyse the patient's Albumin and Bilirubin levels within the algorithm and determine if there are indications of liver disease. However, this approach has the potential pitfall that it analysing such measurements usually requires expert knowledge.

3 The Data

The CCU data is recorded in a Microsoft SQL Server database using a metamodel specifically developed to store the information. Each patient is afforded a *PatientID* upon first encounter with the NHS. Thereafter, each patient visit is recorded as a separate *encounter* with a unique *EncounterID*. This helps to keep track of each unique visit by the patient to the NHS, which may be spread out over time and for different healthcare issues. The NHS database contains a predefined list of *interventions* that may be applied to patients, identified by a unique *InterventionID*. Each intervention further has different *attributes* that can be recorded, each with its own *AttributeID*s. For example, two sample interventions along with their attributes are shown in Figure 1.

In total the database comprises about 250 GB consisting of data of approximately 5,000 patients. Each *encounter* also has a free-form text field associated with it that contains the CCU consultant's notes about the patient. Furthermore, various patient measurements are also recorded such as weight, height, age, initial diagnosis etc. Some patient attributes such as height and weight are recorded

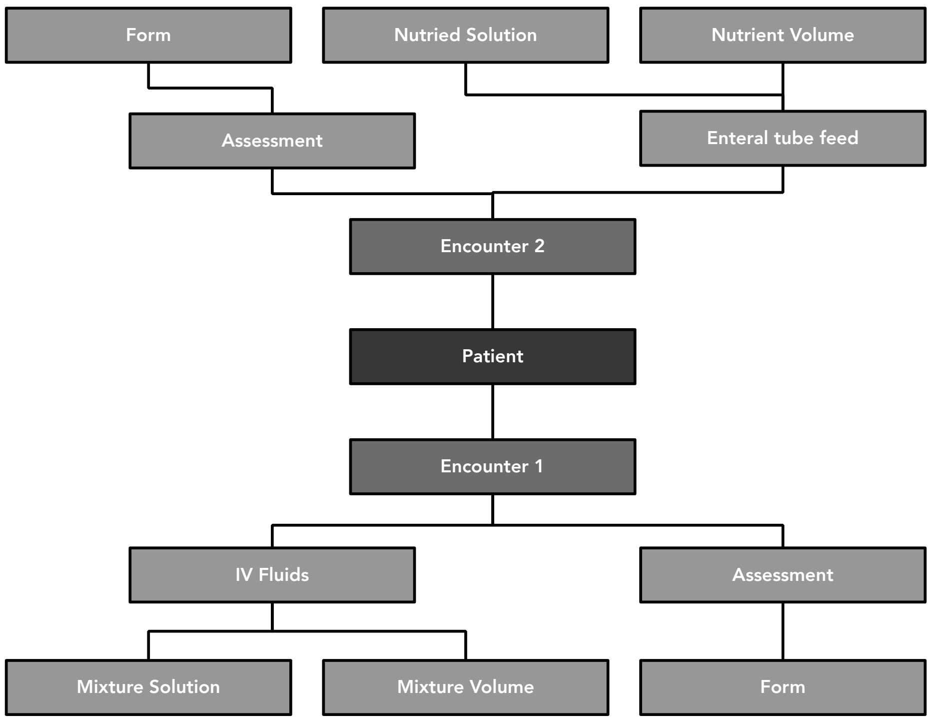

Fig. 1. Simplified NHS Data Model.

only upon admission and / or discharge and additionally as needed. Other attributes, such as blood pressure, heart rate etc are automatically recorded more frequently by the monitoring equipment on the beds. The exact frequency can be set by the CCU consultants such as half hourly, hourly, daily etc. For the purposes of this project, the complexity arises from the fact that much of the information is recorded in the free-form text fields and is not directly recorded anywhere. For example, to determine whether a patient has received little or no nutrition over the past 5 days, we face the same challenge as with determining if the patient has liver disease. We could either aggregate the nutritional intake information recorded in the database, or use natural language processing techniques to extract the same information from the free-form text field. For the former, there are additional issues such as classifying the nutrition as too little or sufficient. Another issue is that the nutritional intake information is not recorded as a single intervention, but in many different interventions. Careful study and foresight is required to accurately extract the necessary information.

4 Future Work

Going forward we need to carefully consider the aforementioned challenges, and devise appropriate strategies and solutions to address them. Various researchers have attempted to solve these problems in various ways. Some researchers have chosen to adopt a statistical approach [6, 11, 2]. Others have attempted to use natural language processing to perform real-time screening of patients based on their history [5, 10]. However, there appears to be a lack of utilisation of other machine-learning techniques for patient screening. We propose that other techniques such as neural networks can also be used to effectively screen patients for dietician referral. A significant challenge in this regard would be synthesis of the required data for input into the learning algorithm. These aspects will be explored further in the coming months.

5 Acknowledgements

We would like to acknowledge the support and help provided by Dr Chris McWilliams, Clinicians, Consultants, and other clinical researchers from the Critical Care Unit at Bristol Royal Infirmary.

References

1. Ghassemi, M., Celi, L.A., Stone, D.J.: State of the Art Review: The Data Revolution in Critical Care. In: Vincent, J.L. (ed.) Annual Update in Intensive Care and Emergency Medicine 2015, vol. 2015, pp. 573–586. Springer International Publishing, Cham (2015)
2. González, C., García-Berrocal, B., Pérez, M., Navajo, J.A., Herraez, O., González-Buitrago, J.M.: Laboratory screening of connective tissue diseases

by a new automated ENA screening assay (EliA Symphony) in clinically defined patients. Clinica Chimica Acta **359**(1-2), 109–114 (Sep 2005). https://doi.org/10.1016/j.cccn.2005.03.042
3. Hall, P., Cash, J.: What is the real function of the liver 'function' tests? Ulster Med J **81**(1), 30–36 (Jan 2012)
4. Hansen, C.C., Dissanaike, S.: Nutrition in the intensive care unit. The Southwest Respiratory and Critical Care Chronicles **3**(10), 10–16 (Mar 2015), https://pulmonarychronicles.com/index.php/pulmonarychronicles/article/view/195
5. He, Q., Veldkamp, B.P., Glas, C.A.W., de Vries, T.: Automated Assessment of Patients' Self-Narratives for Posttraumatic Stress Disorder Screening Using Natural Language Processing and Text Mining. Assessment **24**(2), 157–172 (Mar 2017). https://doi.org/10.1177/1073191115602551
6. Jeong, S., Yang, H., Hwang, H.: Evaluation of an automated connective tissue disease screening assay in Korean patients with systemic rheumatic diseases. PLOS ONE **12**(3), e0173597 (Mar 2017). https://doi.org/10.1371/journal.pone.0173597
7. Lovejoy, C.A., Buch, V., Maruthappu, M.: Artificial intelligence in the intensive care unit. Critical Care **23**(1), 7, s13054–018–2301–9 (Dec 2019). https://doi.org/10.1186/s13054-018-2301-9
8. Mehta, N.M., Bechard, L.J., Cahill, N., Wang, M., Day, A., Duggan, C.P., Heyland, D.K.: Nutritional practices and their relationship to clinical outcomes in critically ill children—An international multicenter cohort study*:. Critical Care Medicine **40**(7), 2204–2211 (Jul 2012). https://doi.org/10.1097/CCM.0b013e31824e18a8
9. Nassar, A.P., Zampieri, F.G., Salluh, J.I., Bozza, F.A., Machado, F.R., Guimarães, H.P., Damiani, L.P., Cavalcanti, A.B.: Organizational factors associated with target sedation on the first 48 h of mechanical ventilation: An analysis of checklist-ICU database. Critical Care **23**(1), 34 (Dec 2019). https://doi.org/10.1186/s13054-019-2323-y
10. Ni, Y., Bermudez, M., Kennebeck, S., Liddy-Hicks, S., Dexheimer, J.: A Real-Time Automated Patient Screening System for Clinical Trials Eligibility in an Emergency Department: Design and Evaluation. JMIR Medical Informatics **7**(3), e14185 (Jul 2019). https://doi.org/10.2196/14185
11. Valentine, N.A., Alhawassi, T.M., Roberts, G.W., Vora, P.P., Stranks, S.N., Doogue, M.P.: Detecting undiagnosed diabetes using glycated haemoglobin: An automated screening test in hospitalised patients. Medical Journal of Australia **194**(4), 160–164 (Feb 2011). https://doi.org/10.5694/j.1326-5377.2011.tb03762.x