MEDINFO 2021: One World, One Health – Global Partnership for Digital Innovation P. Otero et al. (Eds.) © 2022 International Medical Informatics Association (IMIA) and IOS Press. This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0). doi:10.3233/SHTI220087

A Hybrid Approach for Designing Dynamic and Data-Driven Clinical Pathways Point of Care Instruments in Low Resource Settings

Geletaw Sahle Tegenaw^{a,b}, Demisew Amenu^c, Girum Ketema^b, Frank Verbeke^a, Jan Cornelis^a, and Bart Jansen^{a,d}

^aDepartment of Electronics and Informatics ETRO, Vrije Universiteit Brussel (VUB), Pleinlaan 2, 1050 Brussel, Belgium
^bFaculty of Computing, JiT, Jimma University, Ethiopia
^cDepartment of Obstetrics and Gynecology. College of Health Science, Jimma University, Ethiopia
^d iMEC, Kapeldreef 75, 3001 Leuven, Belgium

Abstract

Though a clinical pathway is one of the tools used to guide evidence-based healthcare, promoting the practice of evidencebased decisions on healthcare services is incredibly challenging in low resource settings (LRS). This paper proposed a novel approach for designing an automated and dynamic generation of clinical pathways (CPs) in LRS through a hybrid (knowledge-based and data-driven based) algorithm that works with limited clinical input and can be updated whenever new information is available. Our proposed approach dynamically maps and validate the knowledge-based clinical pathways with the local context and historical evidence to deliver a multi-criteria decision analysis (concordance table) for adjusting or readjusting the order of knowledge-based CPs decision priority. Our finding shows that the developed approach successfully delivered probabilistic-based CPs and found a promising result with Jimma Health Center "pregnancy, childbearing, and family planning" dataset.

Keywords:

Clinical Decision Support Systems, Clinical Pathway, low resource settings.

Introduction

A clinical pathway (CP) is used for maintaining quality care and standardization of the care process through an algorithm, guidelines or evidence [1, 2]. Introducing CP has been shown as a practical and efficient tool for assisting treatment decisions, improving outcomes, and reducing cost [3, 4]. To design a CP, either knowledge-based or data-driven approach have been executed based on the context, principle, and strategy. The knowledge-based CP is designed using paper-based, Information technology (IT)-based, or clinical guideline based (also referred as model-based approaches). The paper-based CP can be delivered from the best experience and local conditions, the IT-based approach automate the paper-based CP and has

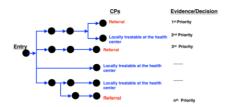


Figure 1: Typical structure (flow) of knowledge-based CPs customized from [22]

been successful to overcome limitations of the paper-based approach [5, 6, 7, and 8], and the clinical guidelines (CGs) can be executed by translating CGs into CP. Information extracted from CG is used as an input for pathway development. However, CGs are "*ambiguous, incomplete and inconsistent*" [9, 10]. In summary, CP can be customized to local conditions and reach consensus among professionals. However, automation and translation of CGs cannot introduce a CP, by definition, rather computerized interpretable CGs [7, 8, 9, and 10]. Thus, contextualizing CP is very crucial and it is important to distinguish the difference between computerized CG and CP.

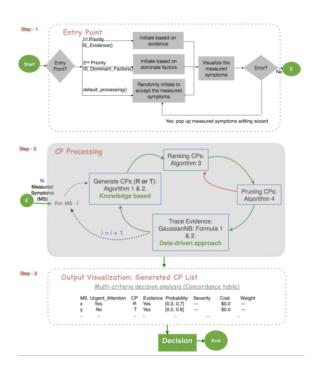


Figure 2: Automated and Dynamic Clinical Pathways: A flowchart for CP entry, Pseudocode for CP processing, and tracing evidence to map and validate knowledge based CPs, and multi-criteria decision output

Whereas the data-driven CP is designed and integrated with the existing electronic health records [11, 12, 13, and 14]. Ontology, probabilistic, and neural network techniques are adopted to enable efficient and adaptive service [15, 16]. However, dealing with ontology-based CP summarization (or presentation) are challenging because of the complex nature of CP variation analysis, accepted standards, and hospital (or healthcare) characteristics. In the case of the probabilistic-based CP [17], when the number of attributes increases, resulting in an exponential number of lookups to deliver recommendations. Though the neural network has shown a promising result for reducing cost and predicting CP variation analysis, learning from small and messy data, and intelligent data acquisition requires further exploration [24, 25, and 26].

However, putting evidence into practice is particularly challenging in low resource settings (LRS), and the gap results in "ineffective treatments that drain the health systems" [17]. The current study highlights two research challenges: (I). Despite the fact that CP has been focused on data-intensive contexts, a significant amount of CP evidence is kept in paper files/charts and circulated in hospital settings [19], and (II). A typical CP guideline-based recommendation, as illustrated in Figure 1, provides referrals and locally treatable pathways based on the observed signs and symptoms. However, a low priority may become a high priority if historical records are traced (considered). Arbitrariness in entry point selection and presenting multi-criteria evidence (e.g. concordance table) for decision were still difficult. Therefore, there is a need to bring a tradeoff mechanism for designing and implementing CP in LRS. The study brought both the knowledge-based and data-driven techniques together to address the existing challenges and implement a link between evidence and practice for promoting the practice of evidence-based care or services. These aims to reduce delay in (i) the decision to seek care assistance by uncovering referral and locally treatable cases, (ii) identifying the decision priority of the referral and locally treatable cases, and (iii) the processing of accurate and real-time information.

Methods

An automated and dynamic clinical pathway was designed for low resource settings. The automated CP contains a range of functions: (I). Initiating entry point for CP, (II). Generating CP for identifying the referral (R) and local treatable (T) cases based on measured symptoms (MS), combinations of symptoms, and indicators (rule sets). The generated CP list was ranked and pruned based on the ranking and pruning parameters respectively. This design provides the mapping of knowledgebased CP using evidence (or historical records) and hence provides visualization of CP, and (III). Displaying a multi-criteria decision output (or concordance table). The concordance table is used for adjusting or re-adjusting the decision priority of generated CP. The overall flow of the design is presented in **Figure 2**. Here below, we presented a detailed description of how automated and dynamic clinical pathways were designed.

2.1 Entry point for CP

An automated input was designed for accepting the measured signs and symptoms. The aim was to reduce arbitrariness in entry point selection. The CP input is flexible and user friendly for validating the measured symptoms and error correcting. The automated input provides a range of choices for initiating the CP such as using evidence from historical records, dominant factors or dynamically initiate based on the signs and symptoms extracted from the clinical guidelines. The detail of an automated CP entry point is presented in **Figure 2**.

2.2 CP Processing

A dynamic CP algorithm was designed for processing the CP using the measured symptoms. The algorithm starts processing CP using the measured symptoms and a combination of measured symptoms as depicted in **Algorithm 1**. For instance; if *"Fever": 'Yes', "BP": '>=140/90', and "Temperature": '<38⁰c'* are the measured symptoms, then **Algorithm 1** returns {

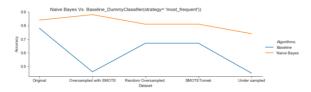
Algorithm 1: Processing Measured Symptoms: The goal is to make ready the measured symptoms and possible combination of measured symptoms for processing (generating) CPs
<pre>Possible.combinations =[]; def possible.combinations =[]; // Search Possible.combination dSymptoms, Possible.combinations(Symptoms, MeasuredSymptoms); for (i = 1; i <= noOFmcasuredSymptoms >1; i = i + 1 } { comb = combinations(MeasuredSymptoms, i); while j in list(comb):): do</pre>

Algorithm 2: CP Processing : Combining Knowledge based with data driven techniques

Possiblecombinations)	edSymptoms = possible.SignsandSymptoms_Measurements(MeasuredSymptoms, // Call the function to process the measured symptoms (i.e. Call Algo 1); G_rulesets_and_indicators.CG_rulesets () //import the CG indicator and rule sets
for ($j = 0$, $j < =$ //comment ; if possibleMas Flag = 7 Generati if the generati if the generation life generation	
Step 5: Prune the CP: Step 6: Trace evidence if Flag == FALSE then	rated.CP_LIST: CP RANKING (Ranking.Parameters): CP PRUNING (Pruning.Parameters) ; for the Generated.PL.IST. Employ Step 4 and 5 while tracing evidence; fed symptoms didn't fulfill the indicator for generating CP 2P_LIST
Algorithm 3: D	ynamic CP in LRS : Ranking CP
	UE) then P ranking based on Evidence ;

Display the ranking difference. If there is any ;
else if IS_choice_Found(TRUE) then
Process the CP ranking based on the choice such as Probability, Severity, or Cost ;
else if IS_manualWeight_Found then
Visualize the CP ranking based on the manual weight refinement;
else
Process accordingly (default processing);
_
Algorithm 4: Clinical Pathway - Pruning
if generated_CP_LIST Is_EMPTY then Go to Algorithm-3 and adjust the Criteria (fall back on CP ranking). Adjust and eliminate one ;
else if too much pruning (TRUE) then
Display eventual warning ;
else
Process the pruning based on the pruning parameters ;
//e.g. CP completeness, CP value (R or T), probability threshold, evidence, and so on); If it FULFILL the endorsed indicator, EXIT;
end

Figure 3: Naive Bayes vs Baseline Classifier Comparison



Then, it generated the clinical pathways based on the output of Algorithm 1 (i.e. measured symptoms and combination of measured symptoms). To generate CPs: (I). Using the first measured symptom, generate all CPs in which the individual symptoms appear, (II). Ranking of CP was performed based on completeness and ranking parameters, (III). Pruning of the dynamic CP list was conducted, (IV). Tracing historical records and the Naive Bayes algorithm was employed for designing data-driven evidence, and (V). Proceed the above steps for the next measured symptom. Overall, the generated CP list was growing and shrinking based on the ranking and pruning parameters. An indicator (also referred as a gold standard and exit criteria) was adopted from the CGs and used for validating the generated CP. If the generated CP already found on the generated list, it will increment the frequency counter. Otherwise, it will add (or append) the generated CP in the generated list of CPs. Algorithm 2 and Figure 2 are depicted to illustrate the detail of CP processing.

2.3 Ranking and Pruning CP

Ranking and pruning were applied to the generated CP as shown in Algorithms **3** and **4**. Ranking of CP provides a range of choices such as ranking the CP based on the evidence, probability, severity, or cost. Whereas pruning was performed to adjust the CP. If the generated CP list is empty, fall back on symptoms ranking and eliminate one, based on evidence. Otherwise, it will process the pruning based on the given parameters (i.e. CP completeness, CP value (not classified, referral or locally treatable), evidence, probability threshold and so on) and exit while fulfilling the indicator. It also displays an eventual warning in case of too much pruning.

2.4 Tracing Evidence

Pregnancy, childbearing, and family planning records from the Jimma Health-Center (Ethiopia) dataset was used for mapping and validating the knowledge-based CP. Data pre-processing including data cleaning, detecting missing and noisy values, and interactive data visualization, data transformation, and data encoding were applied to the health-center dataset. The dataset was a binary CP class (referral and locally treatable cases) with categorical attributes. The dataset contains a total of 63 attributes. NearMiss under-sampling, random oversampling, oversampling with SMOTE (Synthetic Minority Oversampling Technique), and SMOTETomek(combine over-and under-sampling using SMOTE and Tomek techniques) algorithms were applied, experimented, and compared to handle the issue of data imbalance. Then, the Naive Bayes probabilistic technique was implemented and experimented for modeling historical records [23, 24, 27]. Naïve Bayes classifier was employed to explore the probability of features belonging to a CP class (either referral or locally treatable cases) besides the advantage and easyto-use feature. And, the result was compared with the baseline Dummy Classifier with "most_frequent" strategy such as randomly predicts based on the most frequent label in the training set [23]. To train the CP model, a strategy of train test split (test size =0.25) and10-Fold cross-validation methods were

adopted and compared. Accuracy, Matthews correlation coefficient (MCC), receiver operating characteristics (ROC), precision, recall, and F1-score were used to evaluate and measure the performance of the probabilistic CP model and suggest the greatest likelihood for validating the knowledge-based CPs. Finally, an algorithm for tracing evidence was designed. If the measured symptoms and evidence are found on the historical records, then it will load the pre-trained CP model and predict the CP. Otherwise; append the new measured symptoms on the historical records, re-train the model when the number of unseen historical record fulfills the expected threshold, and update the existing CP model.

Results

In this study, an automated CP algorithm was designed for introducing dynamic and data-driven clinical pathways point of care instruments in LRS. A case study was conducted using 719 "pregnancy, childbearing, and family planning" dataset records. Naive Bayes experimentation on the original, oversampled with SMOTE, random oversampled, SMOTETomek, and under-sampled dataset were scored 84, 88, 81, 81, and 74% accuracy, respectively. Whereas the baseline dummy classifier with a strategy of the most frequent was scored 78, 46, 67, 67, and 45% accuracy, respectively. **Figure 3** is depicted to illus-

Table 1: Experimental Result and Metrics

Classifie	Dataset	Status and	Class	Class	Training Strategy						
r Algorith m		Technique s	[Treated/ Referral]						25)	10- Fold Cross Valid ation	
					Accurac y	MCC	Precision	Recall	F1- score	Accur acy	
Naive bayes	Original	Imbalance d and original	532:187 #Instance = 719 #Attributes = 63	Treated	0.84	0.66	0.57	0.97	0.72	82.08	
bayes				Referral			0.99	0.80	0.89		
	Under Near Samplin Algor g	NearMiss	187:187	Treated	0.74	0.54	0.70	0.98	0.82	71.4	
		Aigonuin	#Instance = 374 #Attributes = 63	Referral			0.95	0.48	0.63		
	Random Oversam pling ing algorithm			532:266	Treated	0.81	0.68	0.64	1.00	0.78	84.56
		#Instance = 798 #Attributes = 63	Referral			1	0.73	0.84			
	Oversam pling with SMOTE	SMOTE oversampli ng algorithm	532:532 #Instance = 1064 #Attributes = 63	Treated	0.88	8 0.78	0.81	0.98	0.89	88.6	
				Referral			0.97	0.80	0.88		
	Tomek ove ur san u SN	using	532:266 #Instance = 798 #Attributes = 63	Treated	0.81	0.65	0.64	1	0.78	84.57	
				Referral			1	0.73	0.84		

trate the comparison of the baseline and Naive Bayes classifier accuracy. Information about the distribution of the CP class and the detailed result of the accuracy, MCC, precision, recall, and F1-score using original, under-sampled using NearMiss Algorithm, random oversampled, random oversampled with SMOTE and SMOTETomek datasets are presented in **Table 1**. On our dataset, the accuracy of random oversampling and SMOTETomek was similar. Whereas Matthew's correlation coefficient (MCC) was 68% and 65% respectively. We also

Features Our Hybrid Approach		e-Clinical Pathways [19]	Adaptable Clinical Workflow [20]	Dynamic Clinical Pathways [22]	Methods and systems for clinical process analysis [21]	
Category	-	IEEE Publication	US Patent	US Patent	US Patent	
Low Resource Settings Yes		No	Generic	No	Generic	
Scope	Primary Healthcare			Generic		
Goals	Designing point of care instruments for low resource settings	Creating automated CP data since mostly it's not automated or partially automated (paper-based)	Generating an adaptable clinical workflow (steps) based on a recursive approach	Delivering a system to manage the process using CP programmed as spheres in connected network business rules	Delivering process analysis method for clustering and extracting similar process	
CP Entry (Challenge:arbitrarines s in entry point selection)	Automated and dynamic CP entry based on evidence, dominant factors, or random processing	-	Apply knowledge base and queries to process the reasoning engine	-	-	
Automated	Yes	Yes	Yes	Yes	Yes	
Dynamic (Adaptability)	Yes	-	Yes	Yes	-	
Handling Multiple Clinical Pathways	No	-		Yes	Yes	
Clustering and Extracting Similar CP process	No	No	No	No	Yes	
Inclusion (or exclusion) criteria	Ranking and Pruning Parameters. Probabilities for adjusting the order of CPs.	-	A mix of deductive logic and Bayesian inference for reasoning and Bernoullian value (0-evidence to 1- evidence) used for information expression	A weighted approach using a fuzzy-based on probability to classify weak and strong CPs	Calculating the distances for process clustering	
Outcomes CP evidence tried to reduce the delay by quickly identifying referrals with treated cases.		Automated CP data tried to reduce and impact the length of stay and cost	Generating clinical workflow to increase the quality of care and reduce cost	Helping medical professional to make informed decisions, save costs, and self- learning	Improve the quality of the care process through process clustering and extraction	

Table 2: CP automation and processing methods

made a comparison of train-test-split and 10-fold cross-validation techniques and the result is presented in **Table 1**.

Discussion

The present study was designed to address the existing challenges and implement a link between evidence and practice in LRS. Even though clinical pathways have been developed for decades and concentrated in the data-intensive environments and hospital settings, this paper attempts to deliver a hybrid design approach for introducing automated and data-driven clinical pathways point-of-care instruments in primary healthcare settings. In LRS, paper-based instruments such as hard-copy CGs, patient card-sheets, and point of care charts are the main source of evidence. We have shown an automated and datadriven CPs in LRS, with a well-chosen hybrid design methodology, can provide a multi-criteria decision output (concordance table), ease of access, automated, interactive, and dynamic point of care instrument for assisting the frontline workers and their decision. Naive Bayes based probabilistic techniques were employed for tracing evidence from historical records and a promising result was found. The concordance table helps to adjust (or re-adjust) the decision priority of the CP. More information on the comparison of different CP processing methods and systems are presented in Table 2. However, our comparison is restricted to the scope, goal, input methods, processing techniques, and outcomes. Moreover, the proposed approach provides immediate access to deliver automated and dynamic clinical pathways in LRS by tracing evidence from historical records and dynamically map and validate the generated knowledge-based CPs. However, the case study dataset contains minimum tuple records, imbalanced CP class (i.e. locally treatable and referral cases), and limited to validate and investigate a multi-disease CP generation.

Conclusions

In LRS, the health systems require adaptable service for promoting evidence-based decision. This paper brought a hybrid automated clinical pathway design by combining both knowledge-based and data-driven techniques. Combining the two approaches together may address the existing challenges in LRS and helps to promote the practice of evidence-based decisions and enable automated clinical decision support instruments. Our design has illustrated using 719 "pregnancy, childbearing, and family planning" dataset records from Jimma Health-Center (Ethiopia), and a promising result was found. In the end, a multi-criteria decision table output was presented to arrange (or re-arrange) the decision priority of the CP based on the measured symptoms and local context. Our approach was shown useful for the case study and expected to enhance the documentation of the care process and promote evidence-based health service provisioning. However, exploring large and robust dataset for different parameters and impacts including medication adherence, cost-estimation, and variation analysis as well as investigation of a multi-disease CP generation is warranted.

Acknowledgements

The NASCERE (Network for Advancement of Sustainable Capacity in Education and Research in Ethiopia) program has assisted us in the work to date and will continue to assist us as we move forward with the planned activities.

References

- Kinsman, L., Rotter, T., James, E., Snow, P., & Willis, J. (2010). What is a clinical pathway? Development of a definition to inform the debate. BMC medicine, 8(1), 1-3.
- [2] Deneckere, S., Euwema, M., Van Herck, P., Lodewijckx, C., Panella, M., Sermeus, W., & Vanhaecht, K. (2012). Care pathways lead to better teamwork: results of a systematic review. Social science & medicine, 75(2), 264-268.
- [3] Rotter, T., Kinsman, L., James, E. L., Machotta, A., Gothe, H., Willis, J., ... & Kugler, J. (2010). Clinical pathways: effects on professional practice, patient outcomes, length of stay and hospital costs. Cochrane database of systematic reviews, (3).
- [4] E Abrahams, et al., Clinical pathways: recommendations for putting patients at the center of value-based care. *Clin. Cancer Res.* 23, 4545–4549 (2017).
- [5] S Khodambashi, LA Slaughter, Ø Nytrø, Computer-interpretable clinical guidelines: A review and analysis of evaluation criteria for authoring methods. in *MedInfo*. p. 954 (2015).
- [6] S Chu, Computerised clinical pathway management systems and the implications. *Collegian* 8, 19–24 (2001).
- [7] L DiJerome, The nursing case management computerized system: meeting the challenge of health care delivery through technology. *Comput. nursing* 10, 250–258 (1992).
- [8] Hu, Sheng-Chuan. "Computerized monitoring of emergency department patient flow." The American journal of emergency medicine 11.1 (1993): 8-11.
- K Heiden, Model-based integration of clinical practice guidelines in clinical pathways. in *CAiSE (Doctoral Consortium)*. (Citeseer), (2012).
- [10] B Böckmann, K Heiden, Extracting and transforming clinical guidelines into pathway models for different hospital information systems. *Heal. information science systems* 1, 13 (2013).

- [11] Y Zhang, R Padman, An interactive platform to visualize datadriven clinical pathways for the management of multiple chronic conditions. *Stud. health technology informatics*, 672– 676 (2017).
- [12] JH Bettencourt-Silva, et al., Building data-driven pathways from routinely collected hospital data: a case study on prostate cancer. *JMIR medical informatics* 3, e26 (2015).
- [13] AA Funkner, AN Yakovlev, SV Kovalchuk, Data-driven modeling of clinical pathways using electronic health records. *Procedia computer science* 121, 835–842 (2017).
- [14] Y Zhang, R Padman, N Patel, Paving the cowpath: Learning and visualizing clinical pathways from electronic health record data. *J. biomedical informatics* 58, 186–197 (2015).
- [15] Z Hu, et al., Modeling of clinical pathways based on ontology in 2009 IEEE International Symposium on IT in Medicine & Education. (IEEE), Vol. 1, pp. 1170–1174 (2009).
- [16] D Alexandrou, F Xenikoudakis, G Mentzas, Adaptive clinical pathways with semantic web rules. in *HEALTHINF (2)*. pp. 140–147 (2008).
- [17] R Liu, et al., Pathway-finder: An interactive recommender system for supporting personalized care pathways in 2014 IEEE International Conference on Data Mining Workshop. (IEEE), pp. 1219–1222 (2014).
- [18] Siddiqi, K., & Newell, J. N. (2005). Putting evidence into practice in low-resource settings. 882-882.
- [19] Alahmar, A. D., & Benlamri, R. (2020). SNOMED CT-Based Standardized e-Clinical Pathways for Enabling Big Data Analytics in Healthcare. IEEE Access, 8, 92765-92775.
- [20] Colaert, Dirk, Helen Chen, and Jozef De Roo (2008). "Adaptable Clinical Workflow." U.S. Patent Application No. 12/117,411.
- [21] Bito, Yoshitaka, et al (2009). "Method and system for clinical process analysis." U.S. Patent No. 7,552,062. 23 Jun. 2009.
- [22] Altman, Roy, and Kenneth Altman (2012). "Dynamic clinical pathways." U.S. Patent Application No. 13/152,364.
- [23] "scikit-learn: machine learning in Python scikit-learn 0.23.1 documentation", [online] Available: https://scikitlearn.org/stable/index.html.
- [24] G Du, Z Jiang, X Diao, Y Ye, Y Yao, Variances handling method of clinical pathways based on ts fuzzy neural networks with novel hybrid learning algorithm. *J. medical systems* 36, 1283–1300 (2012).
- [25] MHA Yazid, et al., Clinical pathway variance prediction using artificial neural network for acute decompensated heart failure clinical pathway. *Malays. J. Fundamental Appl. Sci.* 14, 116– 124 (2018).
- [26] M De-Arteaga, W Herlands, DB Neill, A Dubrawski, Machine learning for the developing world. ACM Transactions on Manag. Inf. Syst. (TMIS) 9, 9 (2018).
- [27] Jiang, L., Zhang, H., & Cai, Z. (2008). A novel Bayes model: Hidden naive Bayes. IEEE Transactions on knowledge and data engineering, 21(10), 1361-1371.

Address for correspondence

Geletaw Sahle Tegenaw, Department of Electronics and Informatics ETRO, Vrije Universiteit Brussel (VUB), Pleinlaan 2, 1050 Brussel, Belgium, And Faculty of Computing, JiT, Jimma University, Ethiopia. Email: gsahle@etrovub.be or gelapril1985@gmail.com