Original Article

Agreement of BDMS Utilization Review Technology Version 2 (BURT 2) in Comparison with Utilization Management (UM) Physicians' Opinions to Assess Appropriateness of Continuation of Hospital Stay

Piemchok Banomyong ; Noppadol Phengpinit ; Saowarot Iamsamang; Kusuma Phetchunsakul ; Jinhatha Panyasorn

Piemchok Banomyong

Corporate Utilization Management and Insurance Services, Bangkok Dusit Medical Services, Bangkok, Thailand

*Address Correspondence to author:
Piemchok Banomyong, MD
Corporate Utilization Management and Insurance Services, Bangkok Dusit Medical Services,
2 Soi Soonvijai 7, New Petchburi Rd.,
Bangkok 10310, Thailand
email: piemchok.ba@bdms.co.th

Received: March 16, 2022
Revision received: July 15, 2022
Accepted after revision: September 02, 2022
BKK Med J 2022;18(2): 87-101.
DOI: 10.31524/bkkmedj.2022.21.004
www.bangkokmedjournal.com

Abstract

OBJECTIVES: This study aims to find the agreement between the Bangkok Dusit Medical Services (BDMS) Utilization Review Technology Version 2 (BURT2), an artificial intelligence (AI) driven application, and Utilization Management (UM) physicians' opinions to assess appropriateness of continuation of hospital stay.

MATERIALS AND METHODS: This retrospective study gathered de-identified patients' data from the Health Information System (HIS) of a network hospital of BDMS. The study included patients' data admitted in December 2021. A sample size, calculated from all data, was 274 cases. Inclusion criteria were patients with age more than 3 months and length of stay (LoS) of not more than 7 days (LoS $\leq$ 7 days). All data were processed by BURT 2 to predict the appropriateness of continuation of hospital stay of patients at each admission day. BURT 2 is an AI application specially developed to classify admitted cases on appropriateness for continuation of hospital stay. The application employed convolution neural network (CNN) and natural language processing (NLP) techniques on top of a rule-based algorithm, similar to its predecessor BURT 1. Outputs from BURT 2 were compared with UM Physicians' opinions. BURT 2 was trained until the agreement or accuracy reached 90%.

RESULTS: Among 274 cases, of which 45.3% were male, 53.3% were diagnosed as simple diseases, the majority (42.7%) received services at Internal Medicine Unit. Almost all of cases (95.3%) stayed in hospital for less than four days. The comparison between BURT 2 outputs and UM Physicians' opinions on the appropriateness of continuation of hospital stay in 274 cases showed an agreement of 96%, with 95% sensitivity, 96% specificity, 95% positive predictive value (PPV) and 97% negative predictive value (NPV).

CONCLUSION: BURT 2 had adequate agreement for predicting an appropriateness of continuation of hospital stay. It enabled an initial screening of appropriate continuation of hospital stay, increasing UM nurse work effectiveness, reducing an inappropriate continuation of hospital stay and reducing medical expenses from an inappropriate admission.

Keywords: Utilization Management (UM), Continuation of hospital stay, Artificial Intelligence (AI), BDMS Utilization Review Technology (BURT)

UM is a part of healthcare system that deals with the medical utilization during hospitalization and the medical coverage from insurance companies or third-party payers. A UM Nurse is a nurse who is responsible for reviewing medical records in order to monitor reasonableness, necessity and appropriateness of medical services or admission. The method is carried out by the concurrent review and the retrospective review. This method requires good HIS and completed medical records with systematic reviewing process for ensuring the right decision making. The reviewing process includes a UM physician who is responsible for giving advice and the UM Committee that is responsible for monitoring and determining UM indicators to

standardize the UM works. The examples of indicators were Appropriateness of Admission – Intervention – Consultation – LoS, Utilization Rate, Hospital Day per 1,000 admissions, Average Cost per Case, Average LoS, and etc.[1]

UM is one of the techniques for reducing unnecessary health care requirements, controlling the cost of overutilization that may be a result of misconduct of practitioners or misbehavior of patients and preventing inappropriate admission[2]. Thus, the review and analysis of medical utilization in patients is highly important. It also requires the patients' clinical symptoms review including physical examination results, laboratory results, x-ray results, chief complaint, treatment procedures that were recorded in HIS as well as all medical records that were currently recorded in the electronic medical record (EMR) form. The difficulty and lengthy consumption of the review works possibly causes some human errors in terms of decision making. Recently, a UM nurse of Bangkok Hospital Headquarters routinely reviewed 50 cases a day, varying on the complication of patients' conditions. An experience of each individual UM nurse was another factor that affected the time consumption for utilization review.

This article focusses on the appropriateness of continuation of hospital stay based on physician opinion, assessed by symptoms, physical examination, investigations, and required hospital facilities, for example, the patient has difficulty breathing needs oxygen support in hospital, etc.

This study on the development of BDMS BURT, an application that the investigator team had initiated for analyzing a medical utilization appropriateness, aimed to support the decision of UM nurses when finding an inappropriate treatment. The BURT 2 was upgraded from BURT 1.1. (BURT 1.1 assesses the appropriateness of Admission, while BURT 2 adds a function to assess the appropriateness of Continued Hospital Stay.) The BURT 1.1 applied the platforms of CNN and NPL for its processing process.. These platforms increased their capabilities in terms of wording or meaning classification, latent content interpretation and context decoding to help reduce misevaluation. Also, AI technology was used to process the data into algorithm and to convert written language or natural language into a dummy algorithm for analyzing physician's written language, as seen in Figure 1.

The development of BURT 1.1 indicated that the overall accuracy of appropriate hospital admission was 86.0%. Moreover, the accuracy of BURT when compared with the retroactive approval of insurance claim was 99%. Its processing time was only 0.59 seconds, which is a significant time saving when compared with the time spent by a UM nurse, usually 10-15 minutes per case[3]. Recently, BDMS network hospitals were chosen to be the pilot hospitals to employ BURT version 1.1 as a tool for automatic routine screening. It gathered data directly from HIS database and other systems that relate to treatment of patients during hospitalization. BURT 1.1 could detect the medical records that had a high risk of inappropriate admission, reduce time consumption for reviewing inappropriate admission by UM nurses, increase an effectiveness of claim approval from appropriate and complete medical records, reduce unnecessary medical expenses and decrease possible claim rejection from insurance companies. However, the functions of BURT 1.1 did not cover other health dimensions such as an appropriateness of continuation of hospital stay, etc.

Thus, the Principal Investigator developed BURT 2 to increase the scope of patient safety standards, in particular, the appropriate continuation of hospital stays. The Principal Investigator convened meetings with the UM Physician Panel Team to gather and study guidelines related to the appropriate continuation of hospital stay for each disease[4-8] such as the International Association for Ambulatory Surger[9], Infectious Diseases Society of America[10], Inpatient Discharge Criteria for Children[11] as seen in Figure 2. All related guidelines were applied to determine terms and conditions in BURT 2 (Appendix B).

Materials and Methods

This study was a retrospective chart review. The study protocol was approved by BDMS IRB (COA number BHQ-IRB 2021-11-30). Deidentified data were retrieved from HIS with approval from hospital director of a network hospital in BDMS. The data were patients' data receiving services in December 2021.

BURT 2 was AI application specially developed to classify admitted cases on appropriateness for continuation of hospital stay. The application employed CNN to abstract and interpret free-text data in medical records to see if a word or sentence meets certain criteria. The system has a set algorithm based on data from both NLP and rule-based approach. Techniques were added to a rule-based algorithm, similar to its predecessor BURT.[13]

Subjects were randomly selected from patients who were admitted in December 2021, with age > 3 months and LoS ≤ 7 days. The sample size was calculated using the formula of diagnostic study[14], with the following parameters: sensitivity of at least 80%, prevalence of inappropriate hospital admission of 10%, type I error at 5% ($\alpha < 0.05$) and maximum marginal error of 5%. The calculated sample size was 274 cases.

The evaluation of BURT 2 followed the criteria (Figure 3). Each criterion had 1 score in terms of NLP and rule-based algorithm condition. All scores were summarized into a total score (Figure 4). The score was displayed by the BURT 2 Application into one of 3 types: Inappropriate (0 Score), Borderline (1 Score) and Appropriate (≥ 2 Score). The displayed score was used to reduce UM Nurse workloads in case reviewing. The score indicated the prioritization of the cases. The UM Nurse could allocate their time and efforts to focus on the inappropriate, borderline, and appropriate cases, respectively.

For the criteria for scoring UM physician expert Panel, 3 UM physician experts reviewed the same information form hospital information system (HIS). The information was divided into 6 categories (A→F). The score was given based on medical guidelines and divided into 2 types, appropriate and inappropriate.

The total scores from the UM Physician and BURT 2 were compared. During the development of BURT 2, the criteria were fine-tuned to make BURT 2 as accurate as a UM Physicians Expert panel.

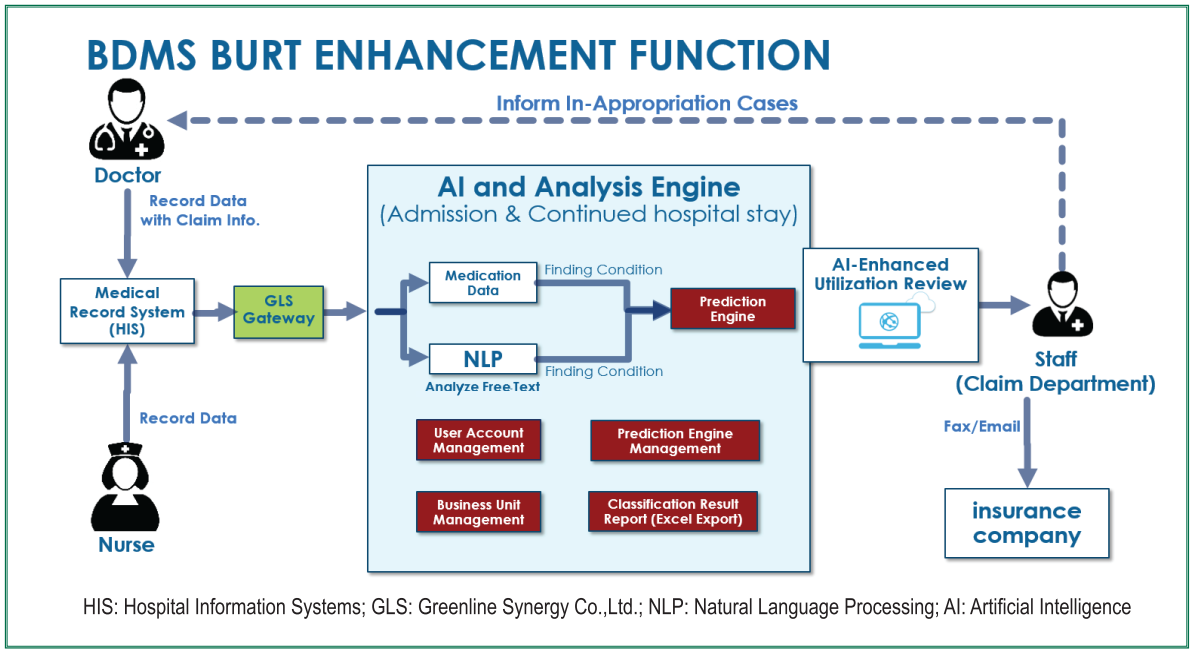


Figure 1: Diagram of BURT Processing System

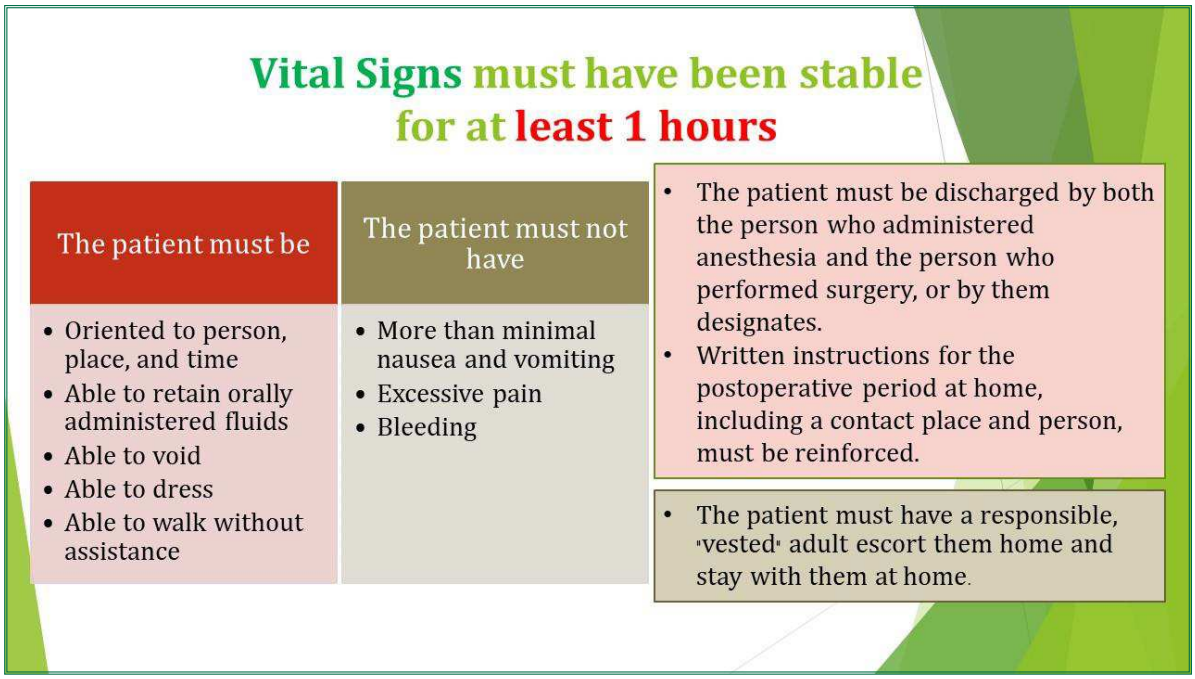


Figure 2: Guidelines for safe discharge after ambulatory surgery[12, 13]

ISSN 2287-0237 (online)/ 2228-9674 (print)

LOS criteria		Classification Technique	Detail	OPD & ER record	Progress note record	T.P.R. record	Pain record	Lab	Imaging	Order sheet	MAR sheet	Order item	Operative note
A History of illness													
A1.1	Neurological/Cardiovascular	NLP	e.g. "หน้ามืด"/"เป็นลม"/ Syncope/ Dizziness		√								
A2.1	Cardiovascular/Lower respiratory problems	NLP	e.g. "หายใจเหนื่อย"/ "เหนื่องง่าย"/Dyspnea		√								
A3.1	Chest pain, Suspected ACS	NLP	e.g. "แน่นหน้าอก"/ Angina/ "ใจสั่น"		√								
A4.3	Unable to eat /Dehydration problems	NLP	e.g. "อาเจียนเป็นเลือด"/ Dysphagia		√								
B Physical Examination													
B1.1	General Appearance	NLP	e.g. "อ่อนแรง"/ fatigue/ ataxia		√								
B3.1	Respiratory Rate	CON	Age < 5 RR <=30 Age >=5 RR <=22			√							
B4	Oxygen Saturation	CON	< 95 %			√							
B5.1	Lower respiratory problems (Abnormal breath sounds), Cardiovascular problems	NLP	e.g. Drooling/Wheezing/ Rhonchi		√								
B6	High Blood Pressure	CON	SBP > 185 or DBP > 120 mmHg			√							
B7.1	Low Blood Pressure	NLP&CON	Orthostatic hypotension SBP < 85 or DBP < 50 mmHg		√	√							
B12.2	Pain score	CON	≥ 5			√	√						
B13.1	Surgical abdomen	NLP	e.g. Rebound tenderness/ Guarding / CVA tender		√								
B14.1	Neurological problems	NLP	e.g. neurological		√								
B15.1	Sign of shock	NLP	e.g. Cold/ Clammy extremities		√								
B22	Temperature	CON	≥ 37.8 c			√							
B24	Count of stool	CON	defecation > 2 times/day			√							
B29	Pulse	CON	0-1 yr (80-140 bpm) 1-2 yr (80-130 bpm) 2-6 yr (75-120 bpm) 6-12 yr (60-110 bpm) >12 yr (40-100 bpm)			√							
C Investigation													
C5	Serum Sodium (Na^+)	CON	< 130 , > 150 mmol/ L					√					
C6	Serum Potassium (K^+)	CON	< 3.1 , > 6 mmol/ L					√					
C7	Bicarbonate (Total CO2)	CON	< 16 , > 34 mmol/ L					√					
C8	Creatinine	CON	> 1.4 mg/ dL					√					
C9	Glucose	CON	< 60 , > 250 mg/ dL					√					
C16	Lactate	CON	> 2 mmol/L					√					
C19	Procalcitonin	CON	> 0.5 ng/ml					√					
D Management													
D1.1	Observation /Monitoring q 4 hrs or more	NLP&CON	e.g. Observe/ Monitor		√					√		√	
D2	Oxygen supplement	CON	order item									√	
D4	Bronchodilator NB at least 2 in 24 hours or m	CON	order item								√	√	
D5.1	IV fluid or IV or IM medication at least 1 in 24 h	CON	order item								√	√	
D7	Blood transfusion or blood components	CON	order item									√	
D10	Troponin- I	CON	2 times in 24 hrs (since the first examination)									√	
D11	Troponin-T	CON	2 times in 24 hrs (since the first examination)									√	
D12	On drain	NLP	on drain		√					√			√
D14.1	Pain killer (Opioid)	CON	Pethidine iv/Morphine iv Fentanyl iv/Tramadol iv (at least 1 dose)									√	
D14.2	Pain killer (NSAID)	CON	on Dynastat ,Ketolac iv (at least 1 dose)									√	
D16	Acyclovir IV drip	CON	at least 1 dose									√	
E Other													
E1.3	Age >= 8 years	CON	criteria	√									
E1.4	Age < 8 years	CON	criteria	√									
F Procedures													
F1	Procedure under Spinal block	CON	criteria										√
F2	Non-minor procedure under General anesthe	CON	criteria										√
F3	Minor procedure under General anesthesia	CON	criteria										√
F4	Cardiac procedures	CON	criteria						√			√	√
F6	Operation time > 90 mins	CON	criteria										√
K Alert													
K1	Simple Disease	CON	criteria	√									
K1.1	non Simple Disease	CON	criteria	√									

*NLP = Natural Language Processing, CON = Value Condition or Data Condition (Rule-based)

Figure 3: Dependent variables and clinical data sources used by BURT predictive algorithm for continuation of hospital stay

Total score = Sum of all criteria scores	
Total score	Predicted result
0	Inappropriate
1	Appropriate (Borderline)
≥ 2	Appropriate

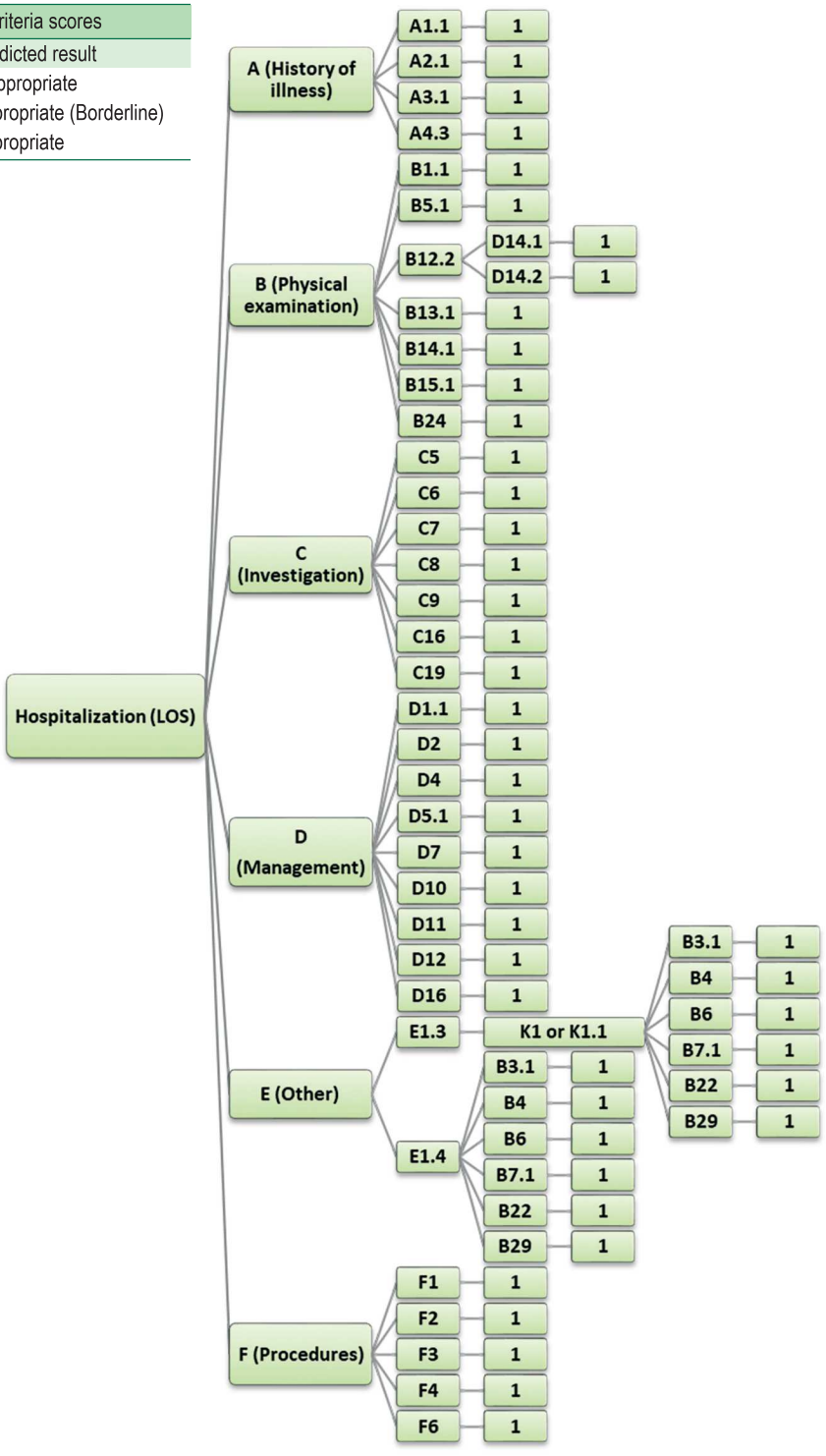


Figure 4: Continuation of hospital stay scoring system implemented in BURT 2 (NLP and rule-based algorithm)

ISSN 2287-0237 (online)/ 2228-9674 (print)

Data collection process

The process of data collection and data processing were performed by hospital IT staff in adherence to standards for protecting patient confidentiality. Recently, Greenline Synergy Co., Ltd. had accessed data under the guideline of ISO/IEC 27001:2013 &ISO 27799 (Information Security Management System). The IT staff gathered data on the inclusion criteria without engaging other data not related to this study.

1. Data related to this study were divided into 6 parts (Figure 3) as follows:
   Part A: History of illness (Data Source: Outpatient department & Emergency room record Admission note)
   Part B: Physical Examination & Vital signs (Data Source: Progress note, Graphic sheet, Pain record)
   Part C: Investigations (Data Source: Laboratory)
   Part D: Medical Care and Management (Data Source: Progress note, Order sheet, MAR sheet, Order items, Operative note)
   Part E: Other; Age (Data Source: Outpatient department & Emergency room record Admission note)
   Part F: Procedures (Data Source: Imaging, Order items, Operative note)
   Part G: Alert (Data Source: Outpatient department & Emergency room record Admission note)
2. Relevant data was uploaded to BURT 2 for processing and evaluating appropriate continuation of hospital stay of patients for each admission day.
3. The data of each patient uploaded to BURT 2 were also distributed to 3 UM physicians with more than 10 years of medical care experience and more than 3 years of UM experience for evaluating and making a decision on appropriate continuation of hospital stay. Each expert was independent from the others. When there were contradictory decisions, the final decision was taken from consensus of 3 experts. The decisions of UM physicians were in line with the gold standard, medical standard of practice.
4. The review from 3 UM physician experts were compared with the outputs from BURT 2 for predicting an appropriate continuation of hospital stay of patient.

Data analysis

The outcomes of this study were measured as follows:

1. Descriptive statistics are used to describe the common characteristics of sample data in this study, including frequency and percentage.
2. An effectiveness of BURT 2 was measured by comparing outputs of BURT 2 with UM Physicians for agreement or accuracy, sensitivity, precision or PPV, Specificity and NPV15.

Results

Among 274 subjects. 45.3% were male. 53.3% were in the group of simple diseases. The majority of sample (42.7%) received services at the Internal Medicine Unit. Almost all of the subjects (95.3%) were admitted with hospital stay of less than 4 days. The details are exhibited in Table 1.

UM Physician Opinions

The decision from the review of 3 UM Physicians in 274 cases showed that the outcomes of 218 cases (80%) were the same, while those of 56 cases, (20%) were not. Medical records of these 56 cases were openly discussed case by case to reach a final decision of appropriate or inappropriate continuation of hospital stay. Then, all decisions were compared with outputs of BURT 2.

Comparing Outputs of BURT 2 and UM Physician opinions

The outcome of the study is shown in Table 2

Table 1: Demographic and clinical characteristics data in this study (n = 274)

Demographic and characteristics	n (%)
Age (years)	
3 months -19	93 (33.9)
20-39	67 (24.5)
40-59	79 (28.8)
≥ 60	35 (12.8)
Sex	
Male	124 (45.3)
Female	150 (54.7)
Disease	
Simple disease	146 (53.3)
Non-simple disease	128 (46.7)
Specialty	
Medicine	117 (42.7)
Pediatric	88 (32.1)
Surgical	23 (8.4)
Cardiology	16 (5.8)
other	30 (11.0)
Length of stay (days)	
1-2	218 (79.6)
3-4	43 (15.7)
5-7	13 (4.7)

*Simple disease was a mild illness without complication or affecting future serious conditions or other diseases. In general, the insurance company usually indicated a group of illnesses. This study referred to simple diseases from the Rational Classification of Simple Disease Cases in BDMS Hospitals using Relative Weight and Case Mixed Index (appendix D).
**Others include departments excluded from the Table such as Ear Nose Throat, Endocrine, Neurology, Obstetric Gynecology.

Table 2: Confusion matrix for appropriate vs inappropriate continuation of hospital stay predicted by BURT 2 prediction output and UM Physician Panel consensus (n = 274)

UMP / BURT	Appropriate	Inappropriate	Total
Appropriate/Borderline	104	6	110
Inappropriate	5	159	164
Total	109	165	274

From Table 2, the overall agreement of appropriate continuation of hospital stay of BURT 2 was 96%, with 95% sensitivity and 96% specificity. Moreover, BURT 2 had shown its precision of 95% PPV, 97% NPV, and 6 cases of false positive cases (false appropriate). After the disagreement was reviewed, it was found that BURT 2 indicated the borderline cases were on the discharge date, which could be interpreted as either appropriate or inappropriate. Three of the 6 cases were interpret as appropriate. The disagreement of another 3 cases were caused by the irrelevant treatment in comparison with the chief complaint (over treatment).

An overall of appropriate continuation of hospital stay interpret by BURT 2 showed 40%. It was similar to 40% that was evaluated by UM physician. An inappropriate continuation of hospital stay interprets by BURT version 2 showed 60% which was similar to 60% evaluated by UM physician.

Table 3 showed the insurance claim approval of 274 cases. The precision or PPV was 100%. The Investigation Team focused on PPV more than sensitivity because the objective of the development of BURT was to detect inappropriate continuation of hospital stay and to manage the case correctly and appropriately before submitting an insurance claim. The table showed that there was one rejected claim from the insurance. A review of the rejected claim found an inappropriate treatment; however, this gap had already been resolved by programing it into the terms and conditions of BURT 2. However, the factor that made the sensitivity low, 40% , was the context of the insurance industry in Thailand. The insurance claim approval was highly flexible. Beside the appropriateness, there were many factors that influenced claim approval.

Table 3: Confusion matrix of BURT prediction output and insurance claim approval

BURT prediction output \ Insurance claim	Appropriate	Inappropriate	Total
Appropriate/Borderline	110	0	110
Inappropriate	163	1	164
Total	273	1	274

Discussion

This study focused on the accuracy of BURT 2 based on the comparison of appropriate continuation of hospital stay between the BURT 2 and the UM physician expert panel. The platform and method of development were similar to BURT 1.1. The structure of BURT 2 is a combination of NLP and rule-based model (condition setting based on Standard Medical Guideline and Expert panel). BURT 2 had been tested until its outcome was similar to UM physician expert panel.

To ensure that BURT 2 could effectively reduce UM nurse workload on concurrent review, the outputs of BURT 2 were also compared with the insurance claim. The outputs excluded the rejected claims that were exempted by the insurance policy (excluded conditions) or business (the coverage without evident medical appropriateness but other mutual benefits). It was expected that the false positive should be less than 5% in terms of medical standards (excluding rejected claim from policy exclusion). The outputs in terms of medical standards were satisfied with 100% precision or PPV. If the prediction of appropriateness of continued hospital stays from BURT 2 was indicated as "appropriate", it was unnecessary for the UM Nurse to review again. UM nurses can focus on inappropriate and borderline cases as a priority, able to manage correctly, appropriately, and in a timely manner. This could reduce workload of UM nurses from insurance claim approval. However, the limitations of BURT 2 was similar to BURT 1.1. They required EMR and computerized physician order entry (CPOE). Its function was also limited if the medical record or physical examination were not complete. In addition, this study was conducted drawing data from one hospital so the variety of data accessed might be insufficient. The training of NLP should be added to cover various contexts and diseases of other hospitals in Thailand.

Recently, the Investigator Team had extended BURT 2 to the pilot hospitals that had implemented BURT 1.1. Furthermore, BURT 1.1 in pilot hospitals was upgraded to BURT version 1.2 (Appendix F). For higher accuracy, BURT 2 was trained in NLP and rule-based by physician expert panel as was BURT 1. Moreover, there was a plan to upgrade BURT 2 to become version 2.1 to expand the screening areas and to cover more investigation and treatment.

Conclusion

The outcome from the development of BURT 2 from NLP and ruled-based model to evaluate the appropriate continuation of hospital stay compared to the outcome from UM physician expert panel was satisfied. BURT 2 reduces the workload of UM nurses in assessment of the appropriateness of continuation of hospital stay. The outputs showed 96% agreement with UM physicians, 95% sensitivity, 96% specificity, 95% precision or PPV and 97% NPV. Its output was complied with insurance claim approval (This study indicated that 237 cases of insurance claims were approved). It supported UM nurses on the claim assessment process and made the management of insurance cases more effective. BURT 2 could build trust with insurance companies and promote medical record completeness. It benefited every related party of UM profession through improved standard of UM, increased appropriate medical treatment, decreased possible complications from prolonged hospital stay and eliminated unnecessary medical expenses.

Acknowledgements

We would like to extend our appreciation to the company taking part in the BURT application development (Softiqal Co., Ltd.), Greenline Synergy staff and Bangkok Health Research Center for your support and suggestions.

References

1. Piravej N, Lohakitsatian A, Piriyaprasarth R, et al. Utilization Management. Bangkok: 2021. (Accessed October 10, 2021 at https://www.si.mahidol.ac.th/th/division/um/admin/download_files/123_48_1toaznc.pdf).
2. Bailit HL, Sennett C. Utilization management as a cost-containment strategy. *Health Care Financ Rev* 1992;1991(Suppl):87-93.
3. Panyasorn J, Banomyong P, Phetchunsakul K, et al. Development of BDMS Utilization Review Technology (BURT): An Artificial Intelligence Tool Using Thai Natural Language Processing to Assess Appropriateness of Hospitalization. *BKK Med J* 2020;16:182-95.
4. Bulut M, Cebicci H, Sigirli D, et al. The comparison of modified early warning score with rapid emergency medicine score: a prospective multicentre observational cohort study on medical and surgical patients presenting to emergency department. *Emerg Med Jl* 2014;31(6):476-81. doi: 10.1136/emermed-2013-202444.
5. Alper E, O'Malley TA, Greenwald J, et al. Hospital discharge and readmission. UpToDate Waltham, MA: UpToDate. 2022. (Accessed October 2, 2021, at https://www.uptodate.com/contents/hospital-discharge-and-readmission?search=hospitaldischarge-and-readmission&source=search_result&selectedTitle=1~150&usage_type=default&display_rank=1).
6. Gaughan J, Gravelle H, Siciliani L. Delayed discharges and hospital type: evidence from the English NHS. *Fiscal Studies* 2017;38(3):495-519. doi: 10.1111/j.1475-5890.2017.12141.
7. Gomez B, Mintegi S, Bressan S, et al. Validation of the "step-by-step" approach in the management of young febrile infants. Pediatrics. 2016;138(2). (Accessed October 2, 2021, at https://doi:10.1542/peds.2015-4381).
8. Queensland government. Practice Guidelines for the Febrile illness Emergency management in children. 2019. Management. (Accessed October 6, 2021,at https://www.childrens.health.qld.gov.au/guideline-febrile-illness-emergency-management-in-children/).
9. International Association for Ambulatory Surgery. Discharge Process and Criteria Following Day Surgery. Bangkok: 2021. (Accessed October 17, 2021, at https://www.iaas-med.com›index.php).
10. Shane AL, Mody RK, Crump JA, et al. Infectious Diseases Society of America clinical practice guidelines for the diagnosis and management of infectious diarrhea. *Clin Infect Dis*2017;65(12):e45-80. doi: 10.1093/cid/cix669.
11. Children's Hospital of Philadelphia. Inpatient Discharge Criteria 2020. (Accessed October 18, 2021, at https://www.iaas-med.com›index.php).
12. Korttila K. Recovery from outpatient anaesthesia: factors affecting outcome. Anaesthesia. 1995;50:22-8. (Accessed October 18, 2021, at https://doi.org/10.1111/j.1365-2044.1995.tb06186.x).
13. Chapter: Clinical Anesthesiology: Anesthetic Management: Ambulatory, Non operating Room, & Office-Based Anesthesia. Anesthesia: Discharge Criteria. [Internet] (Accessed October 18, 2021, https://www.brainkart.com/article/Anesthesia--Discharge-Criteria_27205/).
14. Hajian-Tilaki K. Sample size estimation in diagnostic test studies of biomedical informatics. J Biomed Inforns 2014;48:193-204. doi: 10.1016/j.jbi.2014.02.013.
15. Ostrowski TR, Ostrowski T. The basic four measures and their derivates in dichotomous diagnostic tests. Int J Clin Biostat Biom 2020;6:026. doi: 10.23937/2469-5831/1510026.

Appendix A

Development for Prediction Engine of BURT 2

The steps of medical appropriate continuation of hospital stay prediction by BURT had been designed into 3 layers as follows: (see Figure 4)

1. Concept Layer: This step included the suborders extraction and the processing of AI model. The criteria for extraction could be input in database.
2. Condition Layer: This step had retrieved data from the concept layer and filtered such data by determined terms and conditions.
3. Classification Layer: This step had interpreted all findings into scores. All scores were calculated under the continuation of hospital stay criteria.

The development of BURT 2 had applied NLP as a component of BURT. In the process of physician's written analysis, 77,363 sentences had been trained into the system. An accuracy from the prediction of 23,218 sentences was shown in Table 4. Its accuracy and precision were 98% and 97%, respectively.

Table 4: Confusion matrix for NLP predicted model in BURT 2 (Continuation of hospital stay) (n = 23,218 sentences)

	Predicted Value			
BURT	Actual value		Accuracy	0.98
	Positive	Negative	Precision	0.97
Positive	5,141	171	Recall	0.97
Negative	184	17,722	F1-score	0.97

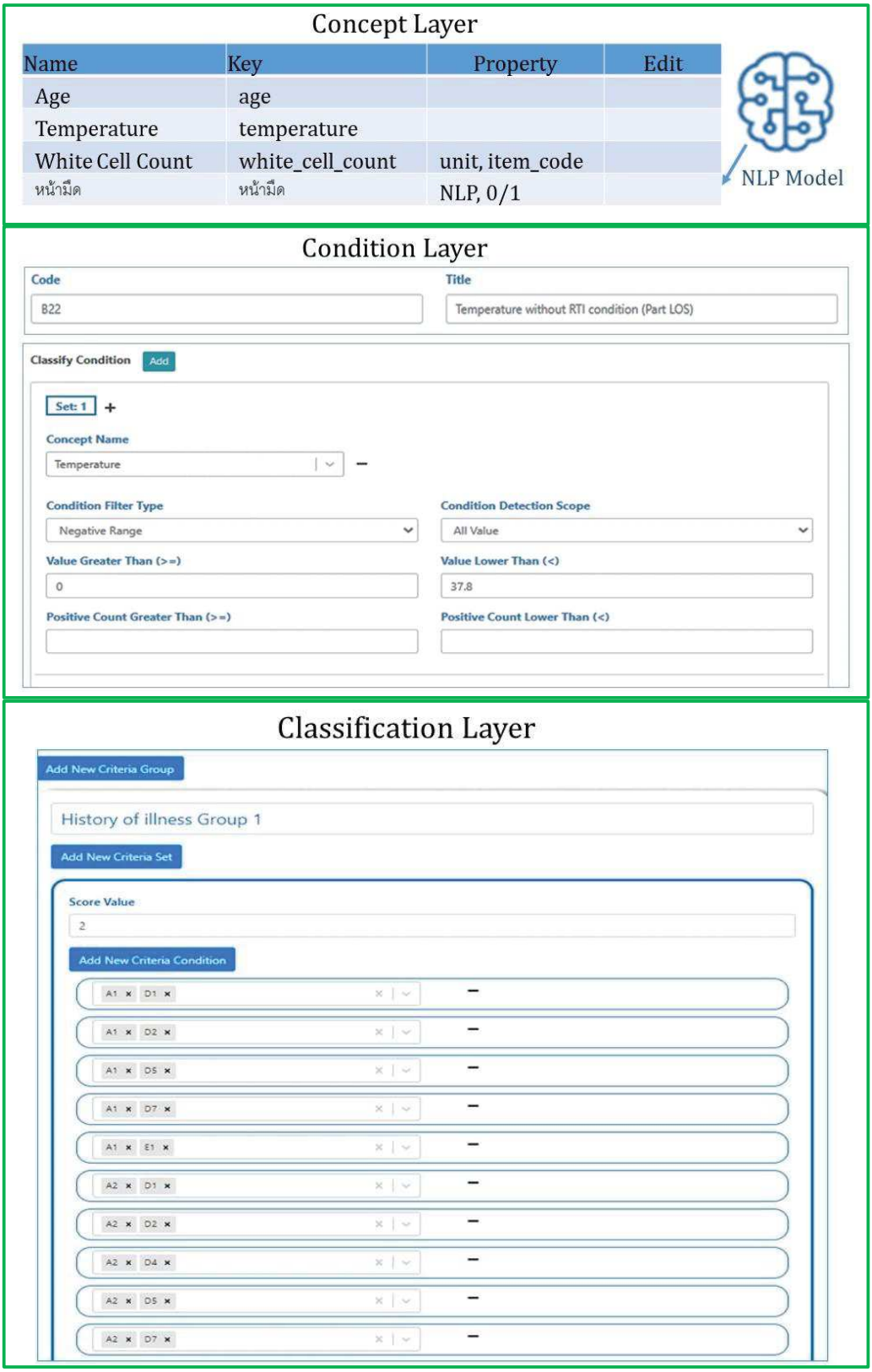


Figure 5: Diagram that presented some example parts to configure the internal relationship in Prediction Engine of BURT 2-part Continuation of hospital stay

ISSN 2287-0237 (online)/ 2228-9674 (print)

Appendix B

Table 5: Examples of Hospitalization appropriateness criteria as implemented in BURT version 2.0

Appropriateness criteria for BURT (Continued hospital stay)		
Variables		Examples of criteria
A History of illness		
A1.1	Neurological/ Cardiovascular	Near syncope, Micturition syncope
A2.1	Cardiovascular/ Lower respiratory problems	Progressive dyspnea, dyspnea at rest
A3.1	Chest pain, Suspected ACS	Unstable angina
A4.3	Unable to eat/ Dehydration problems	Vomiting more than 2 times
B Physical Examination		
B1.1	General Appearance	Angioedema
		Unable to move
		Mental status change
B3.1	Respiratory Rate	RR ≤ 30/min (Age < 5 years)
		RR ≤ 22/min (Age ≥ 5 years)
B4	Oxygen Saturation	< 95 %
B5	Lower respiratory problems (Abnormal breath Wheezing/ Poor air entry/ Stridor/ Rhonchi/ Chest retrac	
B6	High Blood Pressure	SBP > 185 mmHg or
		DBP > 120 mmHg (Age ≥ 9 years)
B7.1	Low Blood Pressure	SBP < 85 mmHg or
		DBP < 50 mmHg (Age ≥ 9 years)
		Orthostatic hypotension (>20 SBP)
B12.2	Pain score	Pain score ≥ 5
B13.1	Surgical abdomen	Rebound tenderness/ Guarding
B14.1	Neurological problems	New abnormal / Focal signs neurological exam
B15.1	Sign of shock	Cold/ Clammy extremities, Capillary refill time (CRT > 2 sec
		Faint pulse
B22	Temperature	Temp ≥ 37.8 °C
B24	Count of stool	> 2 time/day
B29	Pulse	Pulse 80-140 bpm (Age 0-1 years)
		Pulse 80-130 bpm (Age 1-2 years)
		Pulse 75-120 bpm (Age 2-6 years)
		Pulse 60-110 bpm (Age 6-12 years)
		Pulse 40-100 bpm (Age > 12 years)
C Investigation		
C5	Serum Sodium(Na+)	< 130 or > 150 mmol/ L
C6	Serum Potassium(K+)	< 3.1 or > 6 mmol/ L
C7	Bicarbonate (Total CO2)	< 16 or > 34 mmol/ L
C8	Creatinine	> 1.4 mg/ dL
C9	Glucose	< 60 or > 300 mg/ dL with symptomatic
C16	Lactate	> 2 mmol/L
C19	Procalcitonin	> 0.5 ng/ml
D Management		
D1.1	Observation / Monitoring q 4 hrs or more	
D2	Oxygen supplement	
D4	Bronchodilator NB at least 2 in 24 hours or more	
D5.1	IV fluid or IV or IM medication at least 1 in 24 hours or more	
D7	Blood transfusion or blood components	
D10	Troponin- I	> 2 time/day
D11	Troponin-T	> 2 time/day

Appendix C

These figures were the samples of prediction and other details from BURT. The icons of each score had displayed on the screen of BURT 2. The users could use these data for evaluating an appropriate admission and medical appropriate of continuation of hospital stay.

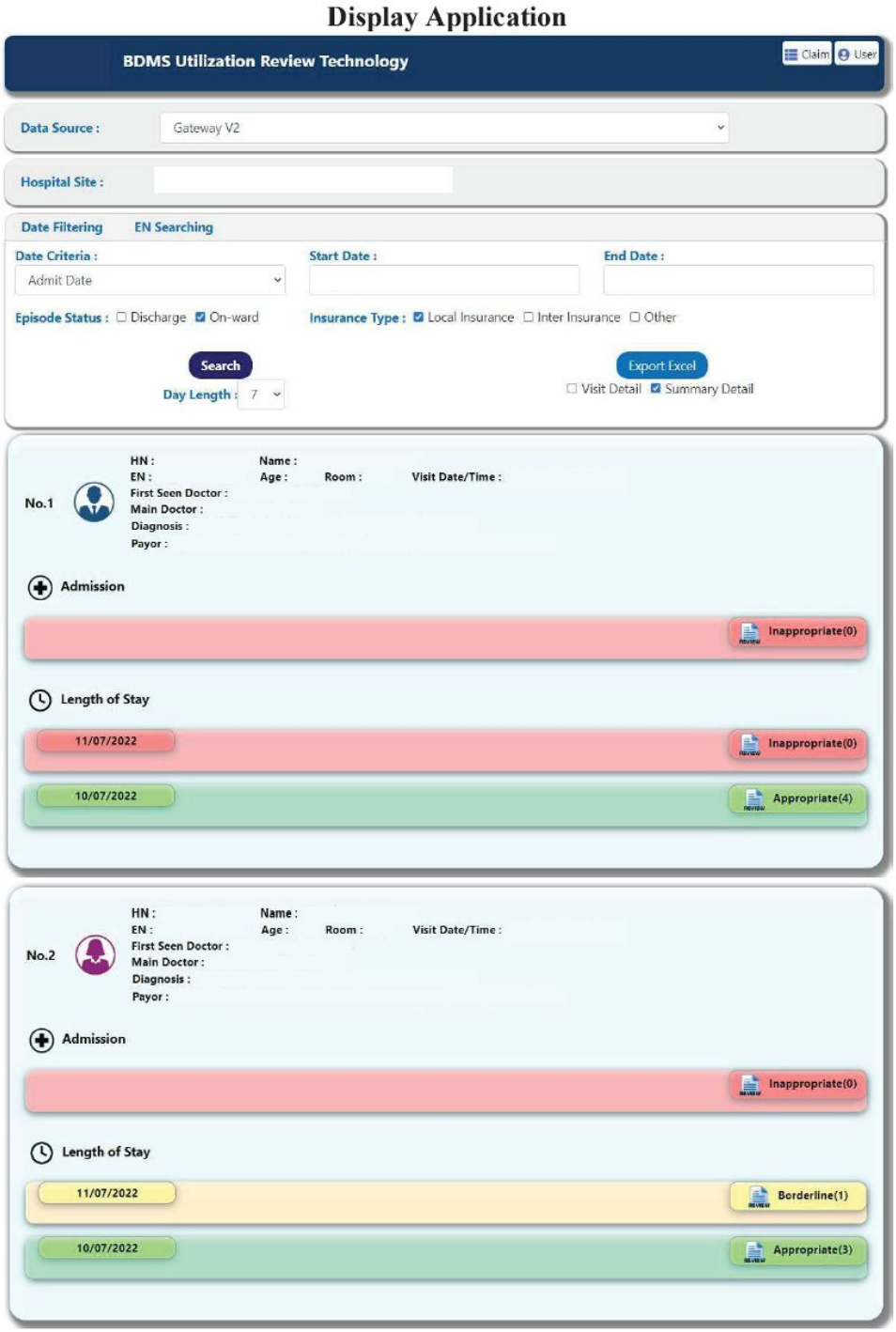


Figure 6: Screen shot of BURT 2 displaying prediction results

ISSN 2287-0237 (online)/ 2228-9674 (print)

Appendix D

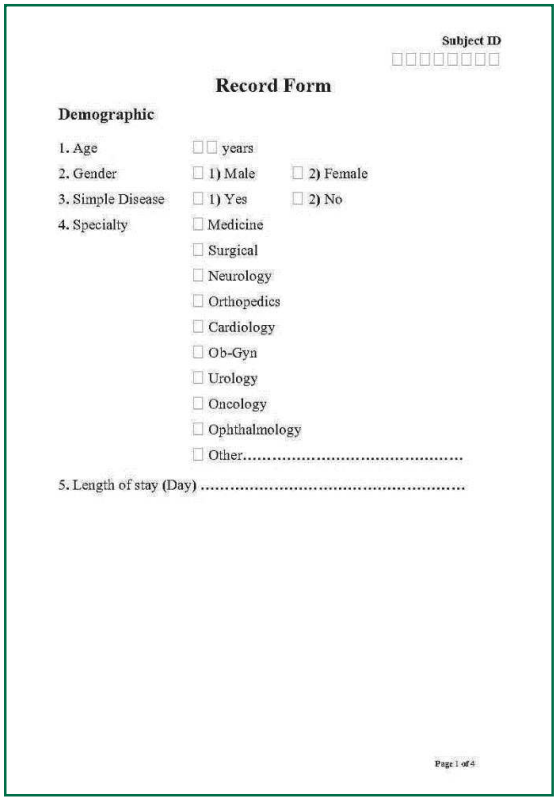
List of Simple diseases

Table 6: Source: "Rational Classification of Simple Disease Cases in BDMS Hospitals using Relative Weight and Case Mixed Index, disseminated in BKK Med J 2019;15(2):130-9."

Simple disease	Relative Weights (RW's)
Vertigo	0.1998
Dizziness	0.1998
Gastroenteritis	0.2225
Diarrhea	0.2225
Acute Gastro-enteritis	0.2225
Gastritis	0.2395
Dyspepsia	0.2395
Dermatitis unspecified	0.2444
Migraine	0.2444
Headache	0.2444
Urticaria	0.2444
Abdominal pain	0.2486
Viral Infection	0.2541
Pharyngitis	0.2591

Simple disease	Relative Weights (RW's)
Influenza	0.2591
URI/URTI	0.2591
Common Cold	0.2591
Acute Sinusitis	0.2591
Acute Tonsillitis	0.2591
Allergic rhinitis unspecified	0.2591
Fever unspecified (child)	0.2667
Syncope	0.2705
Myalgia	0.361
Laryngitis	0.3678
Tracheitis	0.3678
Acute Bronchitis	0.3678
Fever unspecified (adult)	0.3864

Appendix E



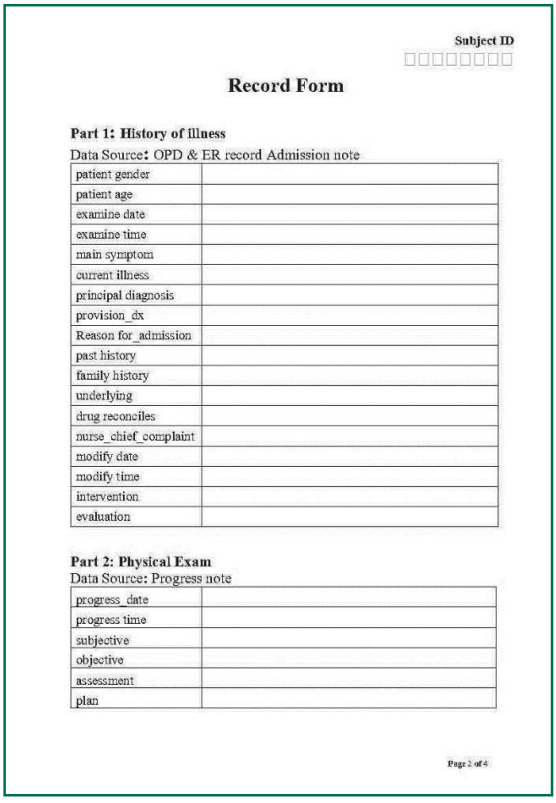
Subject ID □□□□□□□□

Record Form

Demographic

1. Age □□ years
2. Gender □ 1) Male □ 2) Female
3. Simple Disease □ 1) Yes □ 2) No
4. Specialty
   - □ Medicine
   - □ Surgical
   - □ Neurology
   - □ Orthopedics
   - □ Cardiology
   - □ Ob-Gyn
   - □ Urology
   - □ Oncology
   - □ Ophthalmology
   - □ Other.....
5. Length of stay (Day) .....

Page 1 of 4



Subject ID □□□□□□□□

Record Form

Part 1: History of illness

Data Source: OPD & ER record Admission note

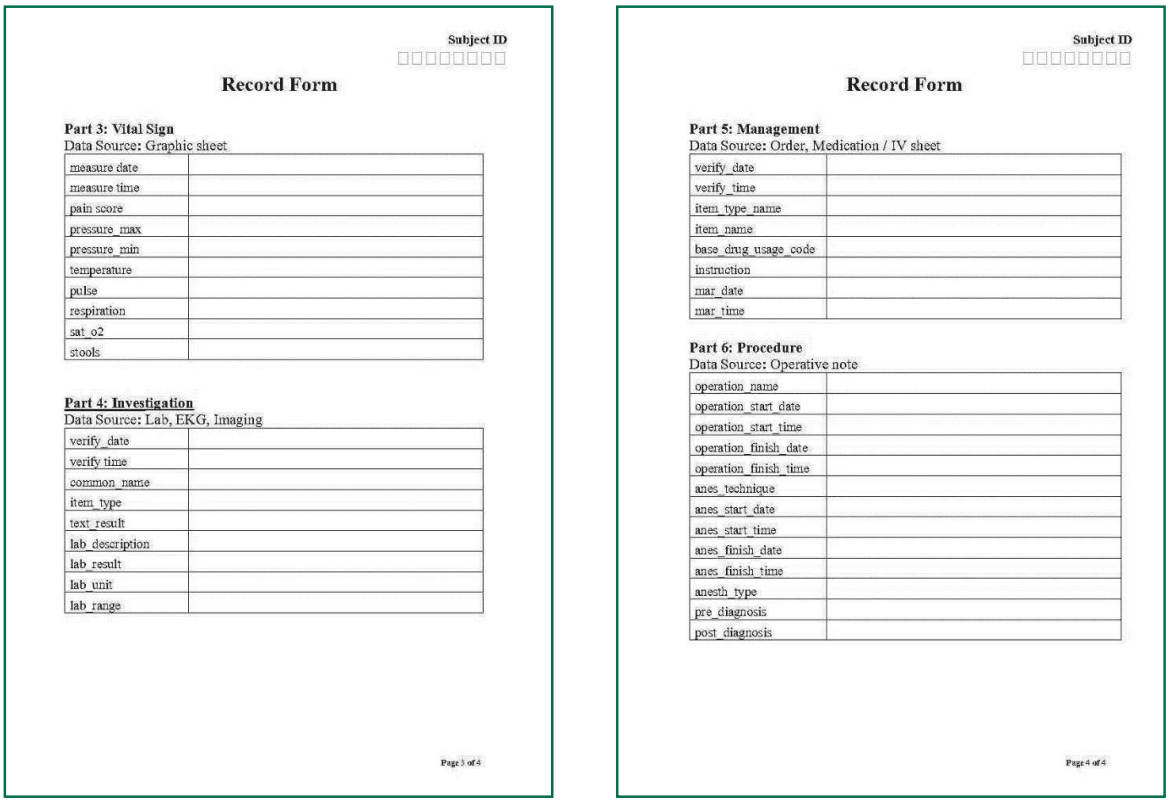
patient gender	
patient age	
examine date	
examine time	
main symptom	
current illness	
principal diagnosis	
provision_dx	
Reason for_admission	
past history	
family history	
underlying	
drug reconciles	
nurse_chief_complaint	
modify date	
modify time	
intervention	
evaluation	

Part 2: Physical Exam

Data Source: Progress note

progress_date	
progress time	
subjective	
objective	
assessment	
plan	

Page 2 of 4



Subject ID □□□□□□□□

Record Form

Part 3: Vital Sign
Data Source: Graphic sheet

measure date	
measure time	
pain score	
pressure_max	
pressure_min	
temperature	
pulse	
respiration	
sat_o2	
stools	

Part 4: Investigation
Data Source: Lab, EKG, Imaging

verify_date	
verify time	
common_name	
item_type	
text_result	
lab_description	
lab_result	
lab_unit	
lab_range	

Page 3 of 4

Subject ID □□□□□□□□

Record Form

Part 5: Management
Data Source: Order, Medication / IV sheet

verify_date	
verify_time	
item_type_name	
item_name	
base_drug_usage_code	
instruction	
mar_date	
mar_time	

Part 6: Procedure
Data Source: Operative note

operation_name	
operation_start_date	
operation_start_time	
operation_finish_date	
operation_finish_time	
anes_technique	
anes_start_date	
anes_start_time	
anes_finish_date	
anes_finish_time	
anesth_type	
pre_diagnosis	
post_diagnosis	

Page 4 of 4

Figure 7: Sample of Continuation of Hospital Stay Form - Case record form (CRF-ST-20211101)

Appendix F

Appendix F presents the progress of the development of BURT 1.1. BURT 1.1 had recently been implemented in some BDMS network hospitals. The development of NLP and Rule-based processes were regularly monitored and improved by the Investigator Team in collaboration with the UM Physician expert panel until it was upgraded to BURT 1.2. We increased keywords for NLP training (from 79 to 119 terms), and increased NLP training and testing dataset (from 77,707 to 196,102 sentences), (Figure 9 and Figure 10).

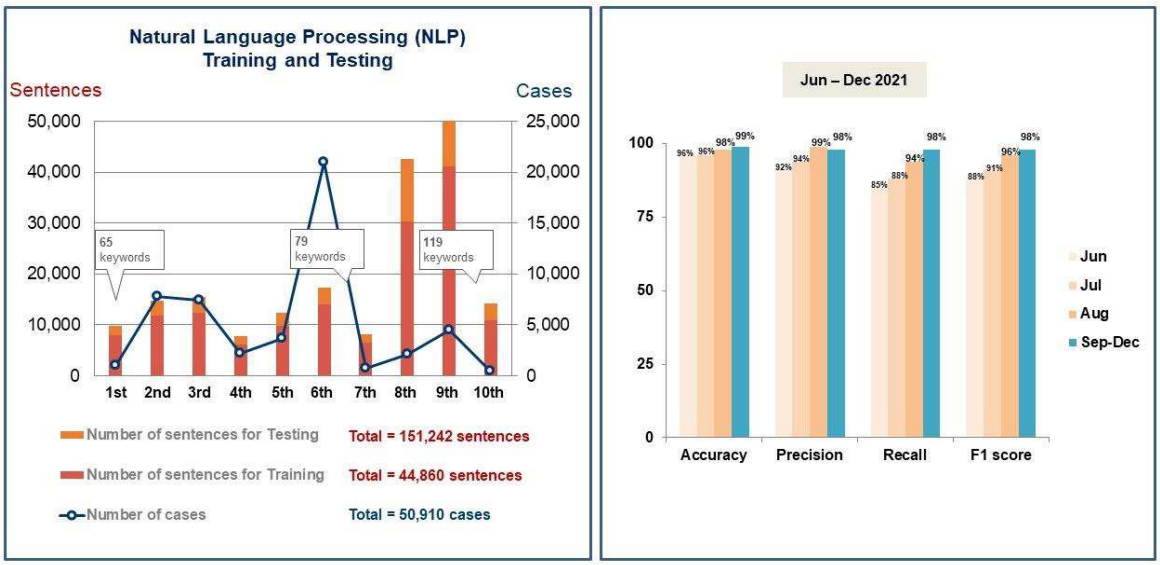


Figure 8: Development of Natural Language Processing (NLP)

ISSN 2287-0237 (online)/ 2228-9674 (print)

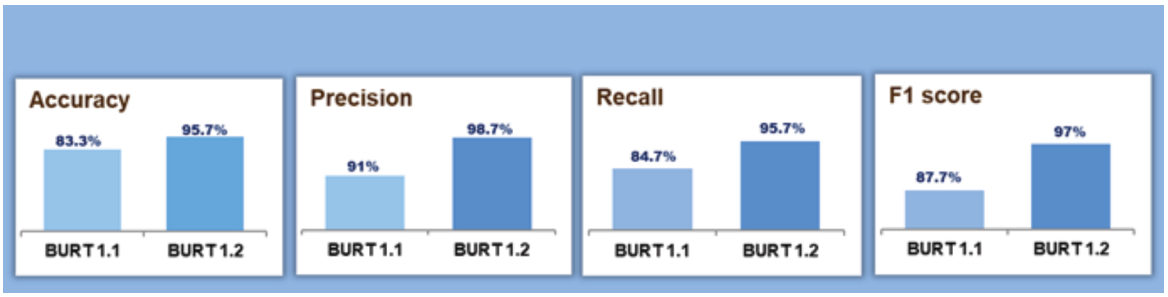


Figure 9: Development of BDMS Utilization Review Technology (BURT 1)

After the implement of BURT 1.1, the satisfaction survey had been conducted. The satisfaction outcome was satisfied (Figure 11). In order to increase satisfaction rate, we drilled down on feedback from users that rated us Fair and Poor. Almost all feedback was about technical problems such as the program frequently encountered errors, the program frequently encountered failure, and the program should be merged with the program for printing concurrently. As for the last feedback, we already had a plan to merge all programs related with insurance into one application to be more user friendly.

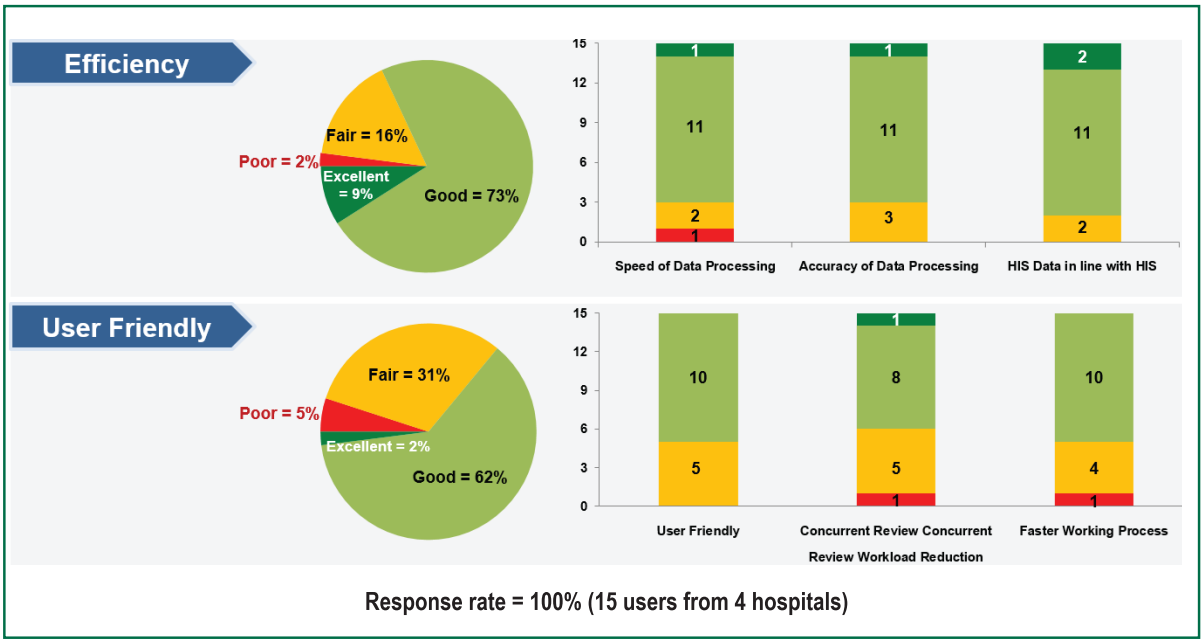


Figure 10: BURT 1.1 User satisfaction survey report