Research on Data – Knowledge-Driven Machine Learning Model for Drug Decision of Cancer Pain

Jian Xiao, Huimin Yu, Lu Zhang, Jingyang Li, Fei Xia, Shuqiao Cheng, Ling Huang

Submitted to: Online Journal of Public Health Informatics
on: May 16, 2024

Table of Contents

Research on Data – Knowledge-Driven Machine Learning Model for Drug Decision of Cancer Pain

Jian Xiao[1, 2, 3] MD; Huimin Yu[1, 2, 3*] MM; Lu Zhang[1, 2, 3*] MD; Jingyang Li[1, 2, 3] MM; Fei Xia[4]; Shuqiao Cheng[1, 2, 3] MD; Ling Huang[1, 2, 3] MM

[1]Department of Pharmacy Xiangya Hospital Central South University Changsha CN
[2]Institute for Rational and Safe Medication Practices National Clinical Research Center for Geriatric Disorders Xiangya Hospital Changsha CN
[3]The Hunan Institute of Practice and Clinical Research Changsha CN
[4]National Supercomputing Center in Changsha Hunan University Changsha CN
[*]these authors contributed equally

Corresponding Author:
Jian Xiao MD
Department of Pharmacy
Xiangya Hospital
Central South University
No. 87 Xiangya Road, Kaifu District
Changsha
CN

Abstract

Background: Drug decision-making is a challenge in the treatment of cancer pain.

Objective: This study aimed to establish machine learning models that can accurately provide drug decisions for cancer pain

Methods: In this study, we built machine learning models based on prior knowledge and clinical data to predict drug decisions for cancer pain treatment. We used 10317 cancer pain treatment medication records from Xiangya Hospital Information System (HIS) and self-developed cancer pain Internet platform (MediHK) to develop decision tree models to classify two kinds of drug decisions for cancer pain treatment and used 1,000 records from Cancer Hospital of Chinese Academy of Medical Sciences for external validation

Results: The two models we developed achieved accuracies of 98.47% and 97.26%, with the AUC of 99.62% and 98.39%, and achieved external verification accuracies of 99.34% and 93.1%, and AUCs of 99.83% and 97.01%, respectively

Conclusions: This study will provide a new model to effectively provide clinicians and pharmacists accurate suggestion, offering a novel approach to untangle the problem of drug decision-making for cancer pain treatment.

(JMIR Preprints 16/05/2024:60600)
DOI: https://doi.org/10.2196/preprints.60600

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Research on Data – Knowledge-Driven Machine Learning Model for Drug Decision of Cancer Pain

Lu Zhang [1,2,3†]; Hui-Min Yu[1,2,3†]; Jing-Yang Li[1,2,3]; Ling Huang[1,2,3]; Shuqiao Cheng*[1,2,3]; Jian Xiao*[1,2,3]

1 Department of Pharmacy, Xiangya Hospital, Central South University, Changsha, China.
2 Institute for Rational and Safe Medication Practices, National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, China.
3 The Hunan Institute of Practice and Clinical Research, Changsha, China.

† *these authors contributed equally*

* *co-corresponding authors*

* Corresponding Author:

Xiao Jian, Prof Dr
Department of Pharmacy
Xiangya Hospital
Central South University
Room 404
No. 87 Xiangya Road, Kaifu District
Changsha, 410008
China
Phone: 86 13973168895
Fax: 86 073184327460
Email: admanoas@163.com

Abstract

Purpose: Drug decision-making is a challenge in the treatment of cancer pain. This study aimed to establish machine learning models that can accurately provide drug decisions for cancer pain.

Methods: In this study, we built machine learning models based on prior knowledge and clinical data to predict drug decisions for cancer pain treatment. We used 10317 cancer pain treatment medication records from Xiangya Hospital Information System (HIS) and self-developed cancer pain Internet platform (MediHK) to develop decision tree models to classify two kinds of drug decisions for cancer pain treatment and used 1,000 records from Cancer Hospital of Chinese Academy of Medical Sciences for external validation.

Results: The two models we developed achieved accuracies of 98.47% and 94.74%, and achieved external verification accuracies of 97.40% and 93.10%, respectively.

Conclusion: This study will provide a new model to effectively provide clinicians and pharmacists accurate suggestion, offering a novel approach to untangle the problem of drug decision-making for cancer pain treatment.

Key words: Cancer pain treatment, Drug therapy, Clinical decision support, Machine learning, Decision tree

Statements and Declarations:

The study was supported by Hunan Provincial Natural Science Foundation of 2022 (2022JJ80114) and The Fundamental Research Funds for the Central Universities of Central South University (2024ZZTS0287).

Author Contributions:

Lu Zhang and Jian Xiao supervised the project and design. Lu Zhang and Hui-Min Yu have accessed and verified the underlying data. All authors participated in data collection.Zhang Lu, Yu Huimin, and Li Jingyang are responsible for the construction of the machine learning model. Lu Zhang and Hui-Min Yu drafted the manuscript. Jian Xiao ang Shu-Qiao Cheng revised the manuscript, and approve the final submission. All of the authors were involved in the acquisition of data and have reviewed and given their approval for the final version of the manuscript to be published.

Introduction

According to statistics from the National Tumor Registration Center in China, it is estimated that there will be 4.82 million new cases of cancer and 3.21 million cancer-related deaths in 2022[1]. Pain is one of the most common symptoms experienced by cancer patients, with a prevalence rate of up to 53%[2], significantly reduces the quality of life for cancer patients and hinders cancer treatment[3]. Therefore, effective management of cancer-related pain is crucial for enhancing patient welfare. With the advancement of modern medicine, the repertoire of cancer pain management has expanded, incorporating methods like interventional therapy, physical therapy, acupuncture therapy, massage, and music therapy[4, 5], thereby broadening the therapeutic landscape. Despite this diversity, pharmacological treatment is the primary method for treating cancer pain. The World Health Organization’s 1986 guidelines for cancer pain management set a benchmark in this regard, offering evidence-based medication strategies to clinicians globally, advocating that healthcare professionals commence with a comprehensive assessment of a patient's pain, subsequently tailoring medication to align with the identified pain severity[6]. And in 2011, the National Health Commission of China, integrating advances in cancer pain treatment from a global perspective and drawing on experiences from Good Pain Management Ward (GPM ward) in China, issued specific guidelines for the diagnosis and treatment of cancer pain tailored to the Chinese context[7]. Despite the support provided by existing guidelines and standards, China continues to face numerous challenges in the actual practice of cancer pain management. According to surveys, ineffective pain management affects 41.3% to 52.9% of cancer patients in China, a figure markedly higher than those reported in North America and Europe, where the rates are 39.1% and 40.3%, respectively[8].

To address this challenge, the implementation of Multi-disciplinary Team (MDT) has become an essential strategy

in managing cancer pain within China[9]. Research has shown that the involvement of pharmacists in MDT-driven outpatient cancer pain management significantly improves pain control and medication adherence in patients[10]. In these MDTs, clinicians are responsible for diagnosis and treatment, while pain specialty pharmacists provide pain assessment, monitoring of drug adverse reactions, evaluating and improving patient medication compliance, and collaborating with clinicians in devising pain medication strategies[10]. However, healthcare facilities at all levels in China, especially at the primary care level, commonly face shortages of clinical pharmacists, particularly in pain management. Moreover, some junior clinicians and pharmacists lack the necessary pharmaceutical knowledge and clinical experience, making it challenging for them to competently address complex decisions in cancer pain treatment.

In this context, the introduction of machine learning (ML) technologies is viewed as a potential solution. Machine learning, a core technology of artificial intelligence, is widely used in the healthcare sector[11]. These models are capable of providing data-driven therapeutic recommendations by analyzing vast datasets of case histories and treatment outcomes[12]. The application of ML in the field of cancer pain is already in progress. However, the majority of studies to date have concentrated on cancer pain prediction[13], analysis of related factors[14], and pain assessment[15]. Despite a few studies addressing medication in cancer pain, these primarily involve predictions of the need for certain opioids or dosage adjustment[16, 17], there remains a lack of research on machine learning models that support medication decision-making in cancer pain management.

In an earlier stage of our study, we reached a consensus on the process and each step of pharmacological care for cancer pain by the Delphi technique, thereby standardizing the pathway for pharmaceutical care of cancer pain[18]. Based on the pathway, we identified features that are important for cancer pain decision-making by extensively consulting expert opinions, employed decision tree algorithms to deeply mine massive and multi-source real data on cancer pain medication, and established models for two core problems in cancer pain medication decision-making: Firstly, we developed the S1 model to determine the appropriate medication decisions for patients who are experiencing pain for the first time and have not previously received any analgesic treatment. This is referred to as the S1 decision. Secondly, for patients who have undergone analgesic treatments but found them either ineffective or intolerable due to side effects, we formulated the S2 model. This is referred to as the S2 decision. The models are designed for patients newly experiencing pain and those with inadequate pain management. By focusing on these two scenarios, our study comprehensively addresses the predominant aspects of decision-making in cancer pain medication.

The decision-making tool is primarily intended for pharmacists. In a clinical setting, patients visit an MDT clinic where clinicians, trained in the use of the tool, conduct consultations. Pharmacists then input the consultation results into

the tool, which provides suggestions for decision-making or modifying treatment. These suggestions are manually reviewed by clinicians and pharmacists, who then reach a consensus and prescribe accordingly. However, the tool remains functional even in clinics with only clinicians and no pharmacists, as the suggested consultation process and elements are fixed. In such cases, clinicians not only conduct the consultation but also assume the pharmacist's role in the MDT clinic: entering pain-related elements from the consultation, manually reviewing the tool's recommendations, and then prescribing. While this approach may increase the workload for clinicians, it is worth noting that the future of clinics services lies in the collaboration between clinicians and pharmacists, each playing their roles in the model of MDT.

The intention behind this decision-making tool is to standardize the decision pathway and assist in decision-making for senior clinicians/pharmacists in MDT clinics, and to provide a learning opportunity for junior care clinicians/pharmacists to enhance their decision-making approach and knowledge base. Overall, while the model is currently in a nascent stage, it should be refined for actual clinical application in the future.

2. Materials and Methods

2.1 Data Source

The data used in this study was from Xiangya Hospital Information System (HIS), which is extensive and comprehensive. It includes basic personal information from admission, physiological indicators, pain evaluations, rational medication plans based on comprehensive evaluations, dynamic medication adjustments, and final treatment outcomes.

2.2 Inclusion and Exclusion Criteria

We extracted the required data from HIS and our internet platform. Due to limitations in HIS interface permissions, we had to extract the data manually. To protect patient privacy, we coded all information that could be used to identify patients. We selected medical records from outpatients who received at least one healthcare between September 2020 and August 2022. The inclusion criteria were that patients who had to have been diagnosed with cancer through pathology or cytology and concurrent pain, aged 18 years or older. We obtained 12,806 medical records, of which 2,832 were in S1 and 9,974 were in S2. Given the large amount of data in the electronic medical record system and the advantages of machine learning, we excluded the data that was missing information and could not be completed (missing value > 20%). Then, we excluded patients with serious complications (e.g., severe hepatic or renal dysfunction). Despite clinical teaching proficiencies present at the home institution for the project, the potential for documentation errors nonetheless existed, possibly stemming from sporadic inaccuracies by medical staff during record-keeping, or inconsistencies within electronic medical records. To mitigate this, we enlisted the expertise of three clinical physicians and two pharmacists from Beihai Second People’s Hospital in Guangxi, The First Affiliated Hospital of Zhengzhou University, Affiliated

Hospital of Guizhou Medical University, and Xiangya Hospital. These experts, each possessing at least five years of clinical experience in cancer pain medication and holding senior professional titles (including associate senior or higher), also hold certificates of training in oncology or cancer pain medication. Following the expert panel's review, erroneous data were eliminated to ensure the accuracy of the decisions.

Figure 1 shows the detailed process of data filtering. Finally, we obtained 10317 cancer pain medication records, with 2397 in S1 and 7920 in S2. The Ethics Committee of Xiangya Hospital at Central South University waived the requirement for informed consent due to the retrospective study design.

2.3 Input variables and Outputs

In this study, we determined the input features for the machine learning model through a multidisciplinary expert panel based on the pharmaceutical care path and knowledge structure of cancer pain determined by the Delphi technique. The panel comprised a clinical pharmacist, an anesthesiologist, an oncologist, and a nurse, all with associate or senior professional titles in oncology. The S1's model included three variables: pain location, pain character, and pain intensity. The S2's model included six variables: pain intensity, pain location, and pain characters, type of breakthrough pain, frequency of breakthrough pain, and types of previous use of analgesics (Table 1). The output of the model was the cancer pain medication decisions recorded in patient medical records by clinicians. Due to the large number of medication types and the explosive growth of decision categories after combination, we simplified the output of the model by classifying medication. For example, hydrocodone sustained-release tablets were classified as strong opioid analgesics and celecoxib capsule as nonsteroidal anti-inflammatory drugs. While judging the correctness of clinical medication decision-making, the expert group also determined the output categories. S1 data obtained 7 output categories and S2 data obtained 18 output categories (Table2 and3 provides output categories and corresponding sample sizes).

2.4 Data Preprocessing

Raw data on cancer pain medication is not suitable for direct modeling without preprocessing. Typical data preprocessing methods include data cleaning, data integration, data transformation, and data coding. Before data processing, we standardized the different names of the original data, such as "burning-like pain" and "scalding pain" as burning pain". Subsequently, we performed data cleaning and transformation.

2.4.1 Data Cleaning

We removed the data with obvious abnormal values and the data that did not correspond to the form of attribute. For duplicate data, we merged useful fields from the duplicate records into one record and then deleted them. This approach allowed us to remove duplicate data and fill in some missing values.

2.4.2 Data Transformation

To reduce attribute values, lower model complexity, and mitigate overfitting risks, we constructed prior features. Pain intensity is the primary indicator for clinicians to evaluate whether patients experience pain and its severity. Our study was conducted at a large tertiary teaching hospital in China, where clinicians use numerical rating scales (NRSs) to evaluate pain intensity, which is homogeneous across clinicians. The NRSs range from 0 to 10, with 0 indicating no pain (this data was not collected in our study as it targets pain patients), 1-3 indicating mild pain, 4-6 indicating moderate pain, and 7-10 indicating severe pain. Therefore, we converted the NRS pain intensity scores into three levels: mild, moderate, and severe [19] (Table 1). Additionally, we combined five pain characters (burning pain, discharge-like pain, stabbing pain, numb-like pain, light touch pain) into neuropathic pain. We also transformed the pain location into four main locations based on clinical experience: head-neck-limbs, joints, viscera, and others. For breakthrough pain frequency, we converted continuous variables into categorical variables: ≥3, 0-3, and 0. And we also classified the types of previous use of analgesics into medication categories.

Finally, we obtained two datasets, one for patients who had not used analgesics before (S1) and one for patients who had used analgesics before (S2). S1 contained 2397 records, and S2 contained 7920 records. S1's model input variables included three types: pain intensity, pain location, and pain characters. The output decision consisted of a combination of nonsteroidal anti-inflammatory drugs, adjuvant analgesics or opioids, with seven categories in total and sample sizes ranging from 193 to 410. S2's model input variables included six types: pain intensity, pain location, and pain characters, type of breakthrough pain, frequency of breakthrough pain, and types of previous use of analgesics. The output decision involved either dynamically adjusting medication based on previous analgesics or making a new medication decision (including dosage adjustment recommendations), with 18 categories and sample sizes ranging from 112 to 818. Table 2 and 3 show the explanations and sample sizes for each output category.

2.5 Decision tree model

In the S1 section of our study, we opted for the decision tree algorithm, specifically employing the classic C4.5 decision tree approach. Throughout this process, the variable chosen at each node is based on the information gain ratio, aiming to maximize the purity of the samples in the child nodes. However, given the susceptibility of decision trees to overfitting, we implemented moderate pruning to eliminate unnecessary branches, thereby enhancing the model's generalizability. In section S2, we constructed models using a range of algorithms, including the Decision Tree (DT) Classifier, Random Forest (RF) Classifier, Ada Boost (AB) Classifier, Support Vector Machine (SVM), Gradient Boosting Classifier (GB), and the Extreme Gradient Boosting (XGB) Classifier. To train and test model performance, each model used 5-fold cross-validation. For each dataset, all data were shuffled and randomly divided into five equally sized subsets. Four subsets were combined as training sets, and the remaining subset was used as a test set. This process

was repeated five times until each subset was used for testing. The training set was used to build models, and the test set was used to adjust hyperparameters and test models' performance. To avoid bias, we conducted five independent runs of 5-fold cross-validation for each model, and the results were averaged over five independent runs. After building the models, we validated S1 and S2 models with external data which was selected from the hospital's medical record system in chronological order of consultation, starting from September 2020. For each model, 1,000 records meeting the inclusion and exclusion criteria were chosen. The model performance was measured by the accuracy (ACC) and area under ROC curve (AUC). Accuracy represents the percentage of correct predictions out of total samples. AUC is a comprehensive indicator that reflects sensitivity and specificity, ranges from 0.5 to 1.0, with values closer to 0.5 indicating lower model accuracy and values closer to 1.0 indicating higher prediction accuracy and performance[20].

3 Results

The S1 decision tree model achieved an average ACC of 98.47% and an AUC of 99.62% through 5-fold cross-validation across five runs (Table 4). The selection of prior features and pruning played a pivotal role in significantly enhancing the overall performance of the model. As shown in Table 5, within the models selected in S2, the Decision Tree model (DT) exhibited a 5-fold cross-validation average accuracy of 94.71%, Random Forest model (RF) 94.20%, Ada Boost model (AB) 87.30%, Support Vector Machine (SVM) 85.61%, Gradient Boosting model (GB) 94.74%, and Extreme Gradient Boosting model (XGB) 94.79%. Among these, the DT, GB, and XGB demonstrated the highest accuracy in classifying the cancer pain dataset. Finally, GB model was selected for S2 decision. The external validation accuracy for the S1 DT model and the S2 GB model were 97.4% and 93.1%, respectively, with AUCs of 99.83% and 97.01%.

In this study, patients seeking analgesic therapy for cancer pain can be divided into two categories: Firstly, cancer pain patients who are experiencing pain for the first time and have not previously received any analgesic treatment. Secondly, those who have undergone analgesic treatments but found them either ineffective or intolerable due to side effects. Based on whether analgesics have been used before, we constructed two decision tree models (S1 and S2). The S1 decision tree model is aimed at patients who have not previously used analgesics, and the S2 decision tree model is aimed at patients who have previously used analgesics. Figure 2 provides visual results of the S1 decision tree model. Taking the left branch as an example, the NRS score, which was the variable with the maximum information gain ratio, was the root node (initial splitting variable) of the decision tree. If the NRS score is≤3, it will enter the left child node for further assessment. If the patient does not exhibit≥2 neuropathic pain characters, they are classified as Class1 (use non-steroidal anti-inflammatory drugs or acetaminophen). If the condition is met, the process proceeds to the next child node, assessing whether the pain location involves joints. If the condition is not met, the classification is class4 (use

anticonvulsants or antidepressants), otherwise, the assessment continues to determine if there are three or more neuropathic pain characteristics, thereby finalizing the classification. Supplementary eMethods(Appendix 1) provides a detailed explanation of all decision-making pathways.

4 Discussion

The decision-making process for cancer pain is a critical aspect of oncology pharmacy services. Many teams in China have attempted to establish a standardized and unified cancer pain decision-making process within their hospitals. However, these processes vary across hospitals, making it challenging to create a universally applicable template. Although some international guidelines offer detailed cancer pain decision-making processes, the clinical and social contexts in other countries may differ significantly from those in China. Our research team convened multidisciplinary experts from tertiary hospitals and utilized the Delphi technique to reach a consensus on each step of the cancer pain pharmaceutical care process[21], thereby standardizing the pathway for cancer pain pharmaceutical care, ensuring the continuity of the cancer pain decision-making process and management, and maximizing the accuracy of patient cancer pain decisions and the safety and efficacy of pain treatment. This provides a foundation for the construction of a decision-making model for cancer pain medication.

Machine learning is widely used to analyze patients' medication status and construct decision models[12]. Guided by the standard pathway for cancer pain pharmaceutical care, we used the decision tree algorithm to deeply mine massive and multi-source real data on cancer pain medication, and establish decision-making models for cancer pain medication. A support vector machine (SVM) classifier was designed to predict the risk of heart failure in participants, with a prediction accuracy of 87.9% [22]. In comparison, the accuracy of the decision models in our study reached 98.47% and 94.74%, respectively, which has the potential to advance our goal of assisting medical personnel in making decisions.

In the clinical field, opaque decision-making systems are not accepted[23]. Simply providing reliable decisions is not sufficient; physicians also need to understand the basis of these decisions. Decision trees, as a type of "white-box" model, possess excellent explanatory power and can graphically display the predictive logic, allowing physicians to easily trace the decision-making path from the root to the leaf nodes and extract clear decision rules[24]. Furthermore, considering that all variables involved in this study are discrete, and decision trees, being classifiers based on information theory, are particularly well-suited for handling discrete feature data. Therefore, we selected decision tree algorithm for S1decision after comprehensive consideration of data characteristics and clinical needs. Comparing the performance of multiple algorithmic models on S2 cancer pain medication decision training set reveals that decision tree algorithms have advantages over other algorithms in such problems. Given the model's accuracy and considerations for subsequent system deployment, the S2 decision compromises by selecting the GB.

Some research refers to expert opinions to narrow down the range of all available data features and further determine the input attributes of machine learning models [25], which is similar to our approach. Feature selection can reduce the dimension of feature space and enhance data quality by removing redundant, irrelevant, or noisy data, thus increasing the comprehensibility of the model and improving its prediction accuracy[26]. We also performed redundancy removal and dimensionality reduction on the data. We combined five pain characters (burning pain, discharge-like pain, stabbing pain, numb-like pain, light touch pain) into neuropathic pain. In addition, we divided pain locations into four main locations based on clinical experience: head-neck-limbs, joints, viscera and others according to the clinical experience to simplify the classification models and increase its generalization ability. The use of prior features significantly improved the overall performance of our models and greatly reduced the sample size required for model training [27].

Using a consensus of preferred treatment decisions reached by a panel of experienced oncologists as the gold standard, researchers compared AI-based clinical decision support system recommendations with treatment decisions made by clinicians[28]. Similarly, we also take expert consensus as the gold standard and invite experts to review the decision data used in modeling and determine the type of model output to ensure the reliability of model training samples.

We attach great importance to the misclassified samples of the validation experiments and have invited an oncologist and a clinical pharmacist to analyze them. The experts believe that the decision results of model training are also correct and interpretable in the real world. For example, in the real world, when a patient's pain intensity is moderate, a physician may choose to discontinue immediate-release weak opioids and increase the dose of sustained-release weak opioids, or replace immediate-release weak opioids or sustained-release weak opioids with sustained-release strong opioids (decision 8), or discontinue immediate-release weak opioids or sustained-release weak opioids and use sustained-release strong opioids (decision 10). This depends on the clinician's strategy to a certain extent, but both decisions are rational. Due to the non-uniqueness of correct decision-making and the subjectivity and randomness of clinicians' decision-making, clinicians may choose another kind of correct decision that is different from that given by the machine learning model. In the future, we will perform multi-labeling on the data and sort the output decisions according to their recommended probability to solve this problem.

Comprehensive and dynamic evaluation of pain is the premise of medication decisions for cancer pain. Evaluation of pain accurately and timely can facilitate the adjustment of the treatment plan and enhance the efficacy of analgesia[29]. Both the NCCN[19] and ESMO[30] emphasized the importance of pain assessment and proposed that pain assessment should

follow the principles of "routine, quantitative, comprehensive and dynamic". However, there are no guidelines to specify the pain assessment process in detail. In addition, the pain assessment process is complicated, so it is still difficult for medical staff to provide comprehensive pain assessment services for patients [11]. Medication selection is a key issue affecting the decision-making of cancer pain management.

Medication therapy is one of the most important strategies in cancer pain treatment [31]. This study is based on the cancer pain pharmaceutical care path and knowledge structure determined by Delphi technology in the early stage. We select six attributes closely related to cancer pain medication decision-making from the path, and build models based on data characteristics. The models include a comprehensive pain assessment, so that the decisions preferred are personalized and comprehensive. Due to the decision tree's interpretability, the model's training results can be derived into several rules. For example, if the type of breakthrough pain is active pain, it will be suggested to choose immediate-release strong opioid; If the type of breakthrough pain is poorly controlled and persistent, it will be suggested to choose sustained-release tablets; If the patient is judged to be suffering from neuropathic pain, the model will recommend anticonvulsants/antidepressants to treat.

Clinical pharmacists, with their professional pharmaceutical knowledge, provide specialized medication advice in multidisciplinary treatment teams (MDT). In cancer pain treatment strategies, the choice of medication is particularly crucial, being one of the decisive factors affecting treatment outcomes. Therefore, clinical pharmacists play a vital role in the decision-making process of cancer pain treatment. However, according to the latest data from the National Medical Products Administration in China (NMPA), there were only 20634 pharmacists in medical institutions in China, by the end of June 2023[32], far fewer than that in developed countries, according to the World Health Organization. This shortfall is especially pronounced in oncology and pain specialties. Moreover, the complexity of decision-making in cancer pain medication is accentuated due to the varying levels of pharmacists' expertise, particularly among junior pharmacists who may lack knowledge and clinical experience. In light of this, we constructed models for two major categories of cancer pain medication decisions to provide clinicians with references of cancer pain medication decisions from pharmaceutical perspectives, thereby assisting experts in confirming and optimizing their decisions, and enriching the knowledge of junior clinicians and junior pharmacists, and help them to learn the experience of experts in the field.

Due to the limited availability of data, the decision-making model for cancer pain medication does not currently incorporate all potentially relevant features. Only attributes deemed important by experts are included, while factors such as medication adherence, basic patient information (e.g., gender and age), vital signs, and socioeconomic status are not considered. Regarding the output category of the model, it is only accurate to the type of medication based on basic facts and medication judgment determined by clinicians, rather than specific medications. Otherwise, the required sample size

would increase exponentially and become unmanageable [33]. This study provides a working template for further precision in decision-making. In the future, we plan to progressively increase the number of classification categories, thereby refining the output decisions in a stepwise manner. Additionally, we aim to incorporate “compliance” a critical attribute. Compared to other features, the measurement of compliance is more subjective. Although compliance information is recorded in the medical record system, it cannot be guaranteed to have been measured by a uniform standard. While some physicians may use compliance scale scores for assessment[34], others might simply make a judgment based on a brief inquiry and their personal experience, hence, an accurate compliance score is unattainable. Moreover, pain management in cancer patients with severe complications is highly complex and requires consideration of numerous factors. Current decision tree models are not yet equipped to handle such complexity; therefore, patients in this group were not included in our study. In addition, there is currently no gold standard for cancer pain medication decision-making, so we can only think that the model we built achieves high accuracy in terms of consistency with expert decision-making.

5 Conclusion

In this study, two kinds of decision models for medication for cancer pain were built, and the evaluation results showed that the accuracy of the S1 and S2 models reached 98.47% and 94.74%, respectively, which could accurately give the decisions for medication of cancer pain. Our model provides a working template for further accurate decision-making, which will enroll the important attribute of medication adherence in the future. The output decisions will be refined step by step by increasing the number of categories. In the future, the model training sample size can be expanded to further refine the decision of cancer pain medication.

Declarations

Ethics approval and consent to participate:

Not applicable

Funding:

The study was supported by Hunan Provincial Natural Science Foundation of 2022 (2022JJ80114) and The Fundamental Research Funds for the Central Universities of Central South University (2024ZZTS0287).

Declaration of Competing Interest:

All authors report no conflict of interest.

Data Sharing Statement:

The data supporting the findings of this study are fully provided within the article.

References

1. Xia C, Dong X, Li H, et al. Cancer statistics in China and United States, 2022: profiles, trends, and determinants. Chin Med J (Engl) 135(5):584-590. http://dx.doi.org/10.1097/cm9.0000000000002108
2. Tawil S, Iskandar K, Salameh P. Pain management in hospitals: patients' satisfaction and related barriers. Pharm Pract (Granada) 16(3):1268. http://dx.doi.org/10.18549/PharmPract.2018.03.1268
3. Kress HG, Aldington D, Alon E, et al. A holistic approach to chronic pain management that involves all stakeholders: change is needed. Curr Med Res Opin 31(9):1743-54. http://dx.doi.org/10.1185/03007995.2015.1072088
4. Virgen CG, Kelkar N, Tran A, et al. Pharmacological management of cancer pain: Novel therapeutics. Biomed Pharmacother 156:113871. http://dx.doi.org/10.1016/j.biopha.2022.113871
5. He Y, Guo X, May BH, et al. Clinical Evidence for Association of Acupuncture and Acupressure With Improved Cancer Pain: A Systematic Review and Meta-Analysis. JAMA Oncol 6(2):271-278. http://dx.doi.org/10.1001/jamaoncol.2019.5233
6. Organization WH. *Cancer pain relief*. 1986.
7. China. NHCotPsRo. Guidelines of Cancer pain Diagnosis and Treatment (2018 edition) (in Chinese). Journal of Clinical Oncology 23(10):937-944.
8. Liu K, Huang H, Zhang L, et al. Effects of a physician- and pharmacist-managed clinic on pain management in cancer patients in China. Basic Clin Pharmacol Toxicol 129(1):36-43. http://dx.doi.org/10.1111/bcpt.13583
9. Zhong X, Zeng X, Zhao L, TaoChen, Min X, He R. Clinicians' knowledge and understanding regarding multidisciplinary treatment implementation: a study in municipal public class III grade A hospitals in Southwest China. BMC Med Educ 23(1):916. http://dx.doi.org/10.1186/s12909-023-04891-0
10. Zhang L, Huang Y, Huang X, et al. Ambulatory care pharmacy practice in China: status and future efforts. Int J Clin Pharm 42(2):321-325. http://dx.doi.org/10.1007/s11096-020-00998-2
11. Fink RM, Gallagher E. Cancer Pain Assessment and Measurement. Semin Oncol Nurs 35(3):229-234. http://dx.doi.org/10.1016/j.soncn.2019.04.003
12. Deo RC. Machine Learning in Medicine. Circulation 132(20):1920-30. http://dx.doi.org/10.1161/circulationaha.115.001593
13. Olling K, Nyeng DW, Wee L. Predicting acute odynophagia during lung cancer radiotherapy using observations derived from patient-centred nursing care. Tech Innov Patient Support Radiat Oncol 5:16-20. http://dx.doi.org/10.1016/j.tipsro.2018.01.002
14. Sipilä R, Kalso E, Lötsch J. Machine-learned identification of psychological subgroups with relation to pain interference in patients after breast cancer treatments. Breast 50:71-80. http://dx.doi.org/10.1016/j.breast.2020.01.042
15. Cascella M, Vitale VN, Mariani F, Iuorio M, Cutugno F. Development of a binary classifier model from extended facial codes toward video-based pain recognition in cancer patients. Scand J Pain 23(4):638-645. http://dx.doi.org/10.1515/sjpain-2023-0011
16. Kumar S, Kesavan R, Sistla SC, et al. Predictive models for fentanyl dose requirement and postoperative pain using clinical and genetic factors in patients undergoing major breast surgery. Pain 164(6):1332-1339. http://dx.doi.org/10.1097/j.pain.0000000000002821
17. Olesen AE, Grønlund D, Gram M, Skorpen F, Drewes AM, Klepstad P. Prediction of opioid dose in cancer pain patients using genetic profiling: not yet an option with support vector machine learning. BMC Res Notes 11(1):78. http://dx.doi.org/10.1186/s13104-018-3194-z
18. Zhang L, Ren XY, Huang HX, et al. Development of the Practice of Pharmaceutical Care for Cancer Pain Management in Outpatient Clinics Using the Delphi Method. Front Pharmacol 13:840560. http://dx.doi.org/10.3389/fphar.2022.840560
19. Network NCC. NCCN clinical practice guidelines in oncology adult cancer pain version 2.2022. 6-7, 2023. Accessed 6-27, 2022. https://www.nccn.org/login?ReturnURL=https://www.nccn.org/professionals/physician_gls/pdf/pain.pdf.
20. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol 5(9):1315-6. http://dx.doi.org/10.1097/JTO.0b013e3181ec173d
21. Zhang L, McLeod HL, Liu KK, et al. Effect of Physician-Pharmacist Participation in the Management of Ambulatory Cancer Pain Through a Digital Health Platform: Randomized Controlled Trial. JMIR Mhealth Uhealth 9(8):e24555. http://dx.doi.org/10.2196/24555
22. Aborokbah MM, Al-Mutairi S, Sangaiah AK, Samuel OW. Adaptive context aware decision computing paradigm

for intensive health care delivery in smart cities-A case analysis. Sustainable Cities and Society 41:919-924.
23. Barredo Arrieta A, Diaz-Rodriguez N, Del Ser J, et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion 58:82-115.
24. Liang M, Wang R, Liang J, et al. Interpretable Inference and Classification of Tissue Types in Histological Colorectal Cancer Slides Based on Ensembles Adaptive Boosting Prototype Tree. IEEE J Biomed Health Inform 27(12):6006-6017. http://dx.doi.org/10.1109/jbhi.2023.3326467
25. Schwartz JM, Moy AJ, Rossetti SC, Elhadad N, Cato KD. Clinician involvement in research on machine learning-based predictive clinical decision support for the hospital setting: A scoping review. J Am Med Inform Assoc 28(3):653-663. http://dx.doi.org/10.1093/jamia/ocaa296
26. Chen JW, Dhahbi J. Lung adenocarcinoma and lung squamous cell carcinoma cancer classification, biomarker identification, and gene expression analysis using overlapping feature selection methods. Sci Rep 11(1):13323. http://dx.doi.org/10.1038/s41598-021-92725-8
27. Yang L, Wu H, Jin X, et al. Study of cardiovascular disease prediction model based on random forest in eastern China. Sci Rep 10(1):5245. http://dx.doi.org/10.1038/s41598-020-62133-5
28. Suwanvecho S, Suwanrusme H, Jirakulaporn T, et al. Comparison of an oncology clinical decision-support system's recommendations with actual treatment decisions. J Am Med Inform Assoc 28(4):832-838. http://dx.doi.org/10.1093/jamia/ocaa334
29. Caraceni A, Shkodra M. Cancer Pain Assessment and Classification. Cancers (Basel) 11(4)http://dx.doi.org/10.3390/cancers11040510
30. Fallon M, Giusti R, Aielli F, et al. Management of cancer pain in adult patients: ESMO Clinical Practice Guidelines. Ann Oncol 29(Suppl 4):iv166-iv191. http://dx.doi.org/10.1093/annonc/mdy152
31. Foley KM, Inturrisi CE. Analgesic drug therapy in cancer pain: principles and practice. Med Clin North Am 71(2):207-32. http://dx.doi.org/10.1016/s0025-7125(16)30866-5
32. NMPA. CCfLPo. The registration status of licensed pharmacists in China in December 2023. 7-1, 2023. Accessed 6-5, 2023.
33. Berisha V, Krantsevich C, Hahn PR, et al. Digital medicine and the curse of dimensionality. NPJ Digit Med 4(1):153. http://dx.doi.org/10.1038/s41746-021-00521-5
34. Thompson K, Kulkarni J, Sergejew AA. Reliability and validity of a new Medication Adherence Rating Scale (MARS) for the psychoses. Schizophr Res 42(3):241-7. http://dx.doi.org/10.1016/s0920-9964(99)00130-9

Table1. Data attributes□description and values

Attributes	Description	Values
Pain location	Location 1(head, neck and limbs), Location 2(joints), Location 3(internal organs)□Location 4(others)	1~4
Pain character	1.Character of neuropathic pain (Stabbing pain, Light touch pain, Numb-like pain, Discharge-like pain, Burning pain), 2. Ache, 3. Jumping pain, 4. Dull pain, 5. Colic, 6. Distending pain, 7. Bearing-down pain, 8. Drill-like pain, 9. Burst-like pain, 10. Tearing pain, 11. Traction pain, 12. Pressing pain, 13. lancinating pain, 14. Innominate pain, 15. Vague pain, 16. Sharp pain,	1~16
Pain intensity	1~3, 4~6, 7~10	1~3
Type of breakthrough pain	1.Activity Pain, 2. End-dose Pain, 3. Persistent Pain	1~3
Frequency of breakthrough pain	≥3, 1~2, 0	1~3
Types of analgesics used previously	1.Non-steroidal/anticonvulsant 2. Non-steroidal/ anticonvulsant + Sustained-release strong opioid 3. Non-steroidal/anticonvulsant+ Sustained-release weak opioid 4. Non-steroidal/anticonvulsant + Immediate-release strong opioid 5. Non-steroidal/anticonvulsant + Immediate-release weak opioid 6. Sustained-release strong opioid+ (non-steroidal/anticonvulsant + strong opioid	1~14

	immediate-release tablets) 7. Sustained-release weak opioid+ (non-steroidal/ anticonvulsant+ Strong opioid immediate-release tablets) 8. Sustained-release weak opioid+ (non-steroidal/anticonvulsant/weak opioid immediate-release tablets) 9. Immediate-release strong opioid + Immediate-release weak opioid 10. Sustained-release strong opioid + Sustained-release weak opioid 11. Sustained-release strong opioid+ Immediate-release weak opioid 12. Sustained-release strong opioid+ Sustained-release strong opioid 13. Immediate-release weak opioid+ Immediate-release weak opioid 14. Immediate-release strong opioid+ Immediate-release strong opioid	

Table2. The rules of S1 decision output data coding

Output	Code	Meaning	Samples
Expert decision-making	1	Use non-steroidal anti-inflammatory drugs or acetaminophen	386
	2	Use strong opioids	389
	3	Use weak opioids	410
	4	Use anticonvulsants or antidepressants	362
	5	Use strong opioids + non-steroidal anti-inflammatory drugs	293
	6	Use strong opioids + anticonvulsants/antidepressants	193
	7	Use weak opioids + anticonvulsants/antidepressants	364

Table3.The rules of S2 decision output data coding

Output	Code	Meaning	Samples
Expert decision-making	1	Stop using immediate-release weak opioids	725
	2	Stop using sustained-release weak opioids	460
	3	Stop using immediate-released weak opioids (recommend to increase the dose of sustained-release strong opioids)	300
	4	Stop using any sustained-release strong opioids	649
	5	Increase the dose of immediate-release weak opioid or stop using the immediate-release weak opioid, switching to sustained-release opioid	235
	6	Stop using the immediate-release weak opioid, and switch to sustained-release opioid	584
	7	Stop using the immediate-release weak opioid, increasing the dose of sustained-release weak opioid or stop using the sustained-release weak opioid, increasing the dose of immediate-release weak opioid or stop using the immediate-release weak opioid and sustained-release weak opioid, switching to strong opioid	469
	8	Stop using the immediate-release strong opioids, and increase the dose of (sustained-release strong opioids or immediate-release strong opioids) by 25 % -50 %	202
	9	Stop using sustained-release weak opioids and immediate-release strong opioids, switching to sustained-release strong opioids, or stop using sustained-release weak opioids, increasing the dose of immediate-release strong opioid	123
	10	Increase the dose of strong sustained-release opioids by 25 % -50 %	625
	11	Increase the dose of sustained-release weak opioids or stop using sustained-release weak opioids, switching to sustained-release strong opioid	479
	12	Stop using the immediate-release weak opioids, add immediate-release strong opioid	228
	13	Stop using the immediate-release strong opioids, and increase the dose of (sustained-release strong opioids or immediate-release strong opioids) by 50 % -100 %	498
	14	Increase the dose of sustained-release strong opioids while increasing the dose of immediate-release strong opioid	632
	15	Stop using the immediate-release weak opioids, and switch to strong opioid	282
	16	Increase the dose of sustained-release strong opioids by 50 % -100 %	818
	17	Add weak opioid	112
	18	Add strong opioid	499

Table 4 Cross-validation results of the decision tree model in the S1 dataset

configuration		Cross-validation											
Prior feature	pruning	1		2		3		4		5		Average	
		ACC[a]	AUC[b]	ACC	AUC	ACC	AUC	ACC	AUC	Acc	AUC	Acc	AUC
		86.36%	92.40%	79.75%	89.33%	85.74%	91.72%	83.06%	90.42%	86.13%	91.70%	84.21%	91.11%
	√	83.68%	96.76%	79.75%	96.82%	80.99%	96.66%	82.02%	96.66%	85.30%	96.51%	82.35%	96.68%
√		97.31%	98.82%	96.07%	98.05%	97.73%	98.91%	96.28%	98.57%	97.72%	98.95%	97.02%	98.66%
√	√	98.55%	99.71%	97.93%	99.18%	97.93%	99.58%	98.97%	99.72%	98.96%	99.90%	98.47%	99.62%

[a]: Cross-validation accuracy of decision tree model

[b]: Area under ROC curve

Table 5 Cross-validation results of the models in the S2 dataset

Model	ACC [a]					
	1	2	3	4	5	mean
DT	95.38%	94.57%	94.36%	93.95%	95.31%	94.71%
RF	94.57%	93.95%	93.95%	93.95%	94.56%	94.20%
AB	85.05%	87.09%	88.03%	91.30%	84.98%	87.30%
SVM	85.53%	86.48%	85.32%	84.57%	86.13%	85.61%
GB	95.38%	94.90%	94.22%	93.88%	95.31%	94.74%
XGB	95.31%	94.97%	94.50%	93.95%	95.24%	94.79%

[a]: Cross-validation accuracy of decision tree model

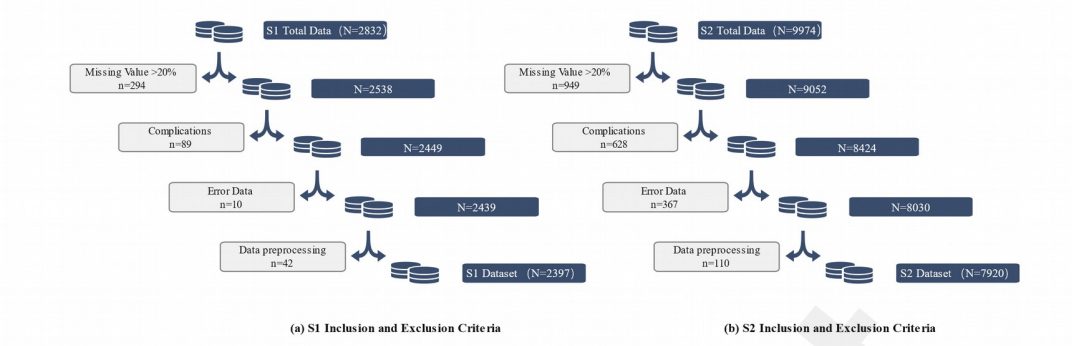


Fig1. Data Exclusion Flowchart

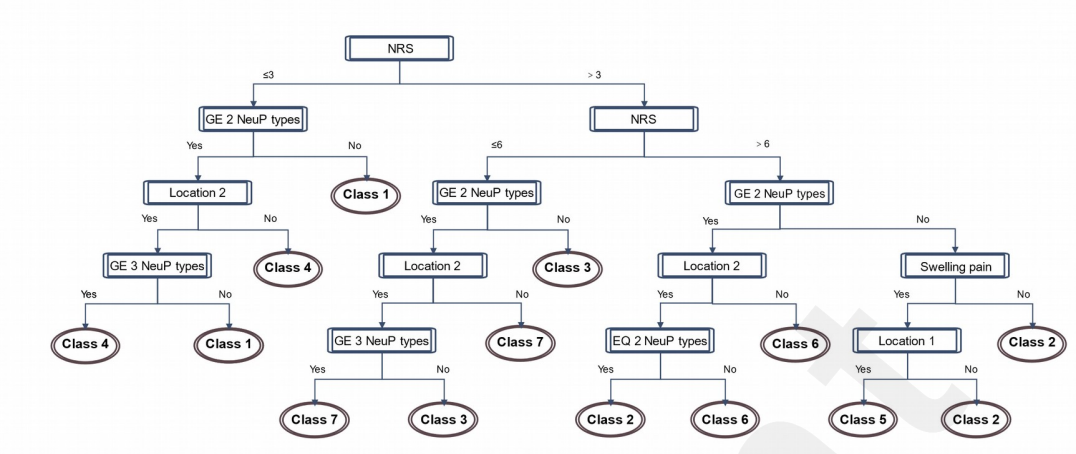
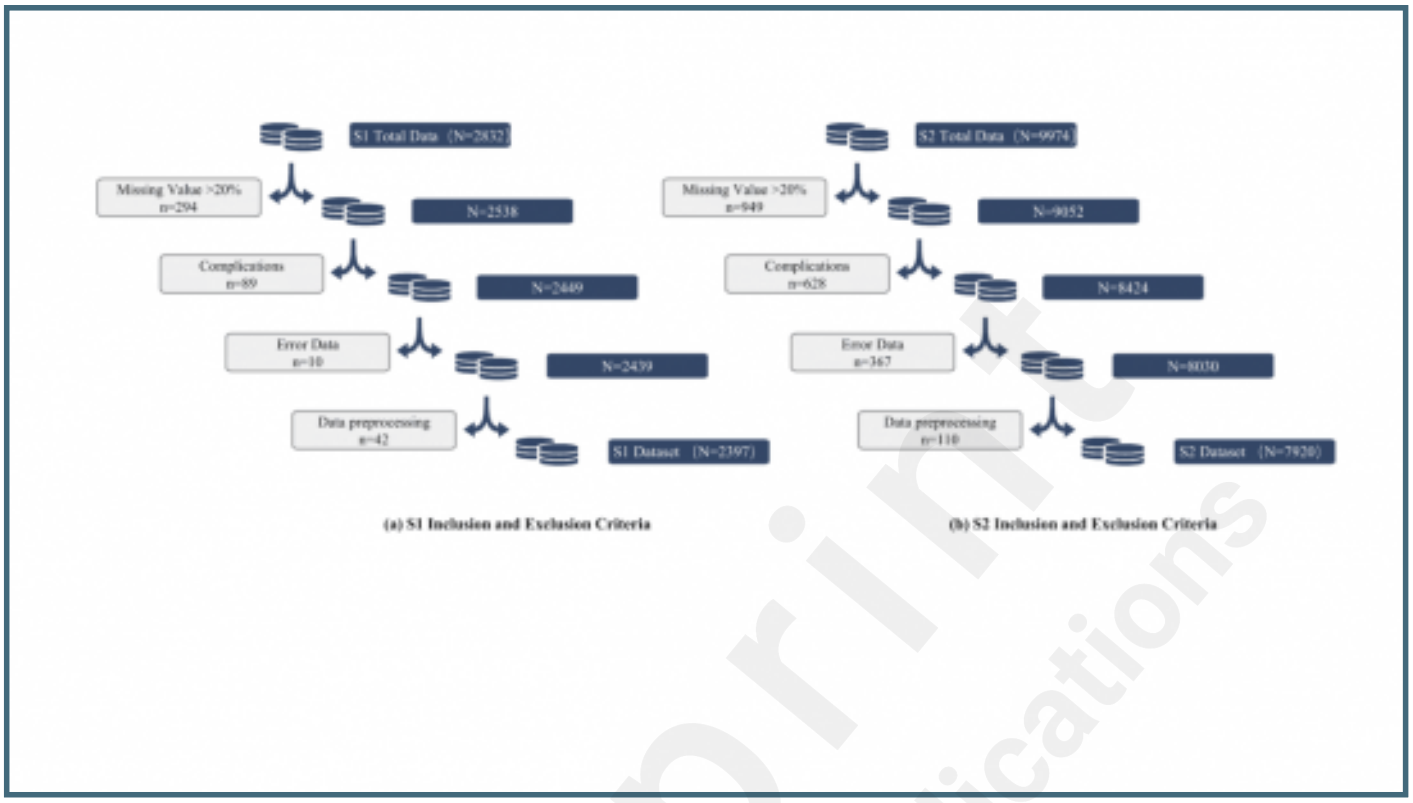


Fig2. Decision tree of S1: NRS: pain intensity; GE 2 NeuP types : whether the pain characters include two or more of the neuropathic pain (Burning pain / Discharge-like pain / Stabbing pain / Numb-like pain / Light touch pain) or not; GE 3 NeuP types : whether the pain characters include three or more of the neuropathic pain (Burning pain / Discharge-like pain / Stabbing pain / Numb-like pain / Light touch pain) or not; EQ 2 NeuP types : whether the pain characters include two of the neuropathic pain (Burning pain / Discharge-like pain / Stabbing pain / Numb-like pain / Light touch pain) or not; Location 1: whether the pain location include any of head, neck and limbs; Location 2: whether the pain location include internal organs

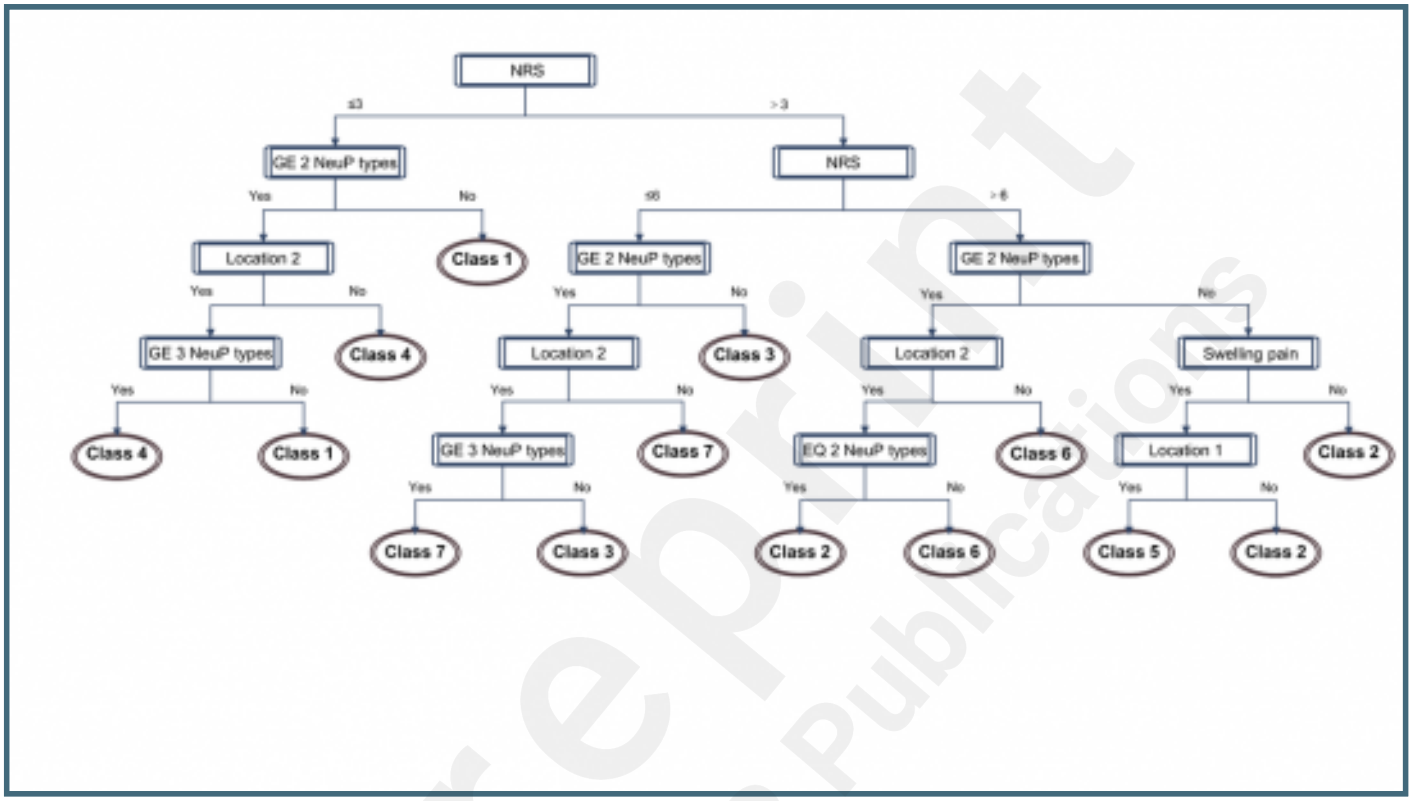
Supplementary Files

Figures

Data exclusion flowchart.



Decision tree of S1: NRS: pain intensity; GE 2 NeuP types : whether the pain characters include two or more of the neuropathic pain (Burning pain / Discharge-like pain / Stabbing pain / Numb-like pain / Light touch pain) or not; GE 3 NeuP types : whether the pain characters include three or more of the neuropathic pain (Burning pain / Discharge-like pain / Stabbing pain / Numb-like pain / Light touch pain) or not; EQ 2 NeuP types : whether the pain characters include two of the neuropathic pain (Burning pain / Discharge-like pain / Stabbing pain / Numb-like pain / Light touch pain) or not; Location 1: whether the pain location include any of head, neck and limbs; Location 2: whether the pain location include internal organs.



Multimedia Appendixes

A detailed explanation of all decision-making pathways.
URL: http://asset.jmir.pub/assets/54f3def779e1f1c75aef51d680a01bf7.docx