Improving Benign Paroxysmal Positional Vertigo Diagnosis

Sida Zhang
Marquette University

Follow this and additional works at: https://epublications.marquette.edu/theses_open

Part of the Engineering Commons

Recommended Citation
https://epublications.marquette.edu/theses_open/777
IMPROVING BENIGN PAROXYSMAL POSITIONAL VERTIGO DIAGNOSIS

by
Sida Zhang, M.A.

Marquette University, 2023

A Thesis submitted to the Faculty of the Graduate School,
Marquette University,
In Partial Fulfillment of the Requirements for
the Degree of Master of Science

Milwaukee, Wisconsin

December 2023
ABSTRACT

IMPROVING BENIGN PAROXYSMAL POSITIONAL VERTIGO DIAGNOSIS

Sida Zhang, M.A.
Marquette University, 2022

Benign Paroxysmal Positional Vertigo (BPPV) is one of the most common causes of dizziness. Especially for people over 45, the risk of BPPV is substantial. On the other hand, BPPV is often misdiagnosed and may require expensive examinations. This thesis introduces a prediction model based on machine learning to quickly, inexpensively, and accurately diagnose BPPV. The thesis starts by introducing BPPV and the statistics of BPPV misdiagnosis. Then, a patient survey is introduced. The patient survey includes 50 BPPV-related questions, which are used as training data for the machine learning model. Logistic Regression, Decision Tree, and Naïve Bayes were compared for machine learning models and their results were discussed. Three machine learning approaches are explored, logistic regression, decision tree, and naïve Bayes with cross validation accuracies of 89.8%, 81.9%, and 75.1%, respectively.
ACKNOWLEDGMENTS

Sida Zhang, M.A.

I would like to thank my parents who supported me and gave me encouragement during the whole time of this work.
# TABLE OF CONTENTS

ABSTRACT ......................................................................................................................... i

ACKNOWLEDGMENTS ...................................................................................................... ii

TABLE OF CONTENTS ...................................................................................................... iii

LIST OF TABLES ................................................................................................................ v

LIST OF FIGURES ............................................................................................................. vi

Chapter 1 Introduction ...................................................................................................... 1
  1.1 Background .................................................................................................................. 2
    1.1.1 Benign Paroxysmal Positional Vertigo ................................................................. 2
    1.1.2 Motivation ............................................................................................................. 3
    1.1.3 Common Symptoms of BPPV .............................................................................. 4
    1.1.4 Diagnostic Methods for BPPV ........................................................................... 5
  1.2 Overview of Thesis ...................................................................................................... 6

Chapter 2 Background for BPPV ...................................................................................... 7
  2.1 Prior Work ................................................................................................................... 7
  2.2 Description of BPPV ................................................................................................... 10
  2.3 Machine Learning Algorithms .................................................................................. 11
    2.3.1 Candidate Algorithms ....................................................................................... 12
    2.3.2 Logistic Regression Classifier ........................................................................... 13
    2.3.3 Naïve Bayes Classifier ..................................................................................... 16
    2.3.4 Decision Tree Classifier .................................................................................... 18
    2.3.5 Cross-validation ............................................................................................... 20

Chapter 3 Survey Process ................................................................................................ 22
  3.1 Aim ............................................................................................................................. 22
  3.2 Structure ..................................................................................................................... 23
    3.2.1 Questions .......................................................................................................... 23
    3.2.2 Scenario ............................................................................................................. 24
  3.3 Improvements of Patient Survey .............................................................................. 25
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3.1 Survey Simplification</td>
<td>26</td>
</tr>
<tr>
<td>3.3.2 In-survey Tutorial</td>
<td>26</td>
</tr>
<tr>
<td>3.3.3 Descriptions of Symptoms</td>
<td>28</td>
</tr>
<tr>
<td>3.4 Survey Questions</td>
<td>29</td>
</tr>
<tr>
<td>Chapter 4 Results</td>
<td>33</td>
</tr>
<tr>
<td>4.1 Data Analysis</td>
<td>33</td>
</tr>
<tr>
<td>4.2 Machine Learning Results</td>
<td>35</td>
</tr>
<tr>
<td>4.3 Discussion</td>
<td>36</td>
</tr>
<tr>
<td>Chapter 5 Conclusion and Future Work</td>
<td>38</td>
</tr>
<tr>
<td>Bibliography</td>
<td>40</td>
</tr>
<tr>
<td>Appendix A</td>
<td>44</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1 Survey Questions .................................................................29

Table 2 Cross-validation result .............................................................35
LIST OF FIGURES

Figure 1 Cumulative incidence of BPPV of different ages ................................................................. 3

Figure 2 Inner ear anatomy (Source: Tetiana Lazunova with permission) ........................................ 11

Figure 3 Logistic distribution ............................................................................................................ 14

Figure 4 Decision Tree Structure .................................................................................................... 18

Figure 5 Age distribution of 40 patients .......................................................................................... 34

Figure 6 Gender of all patients ......................................................................................................... 34

Figure 7 Medical examination results of male and female ............................................................... 35

Figure 8 Home page .......................................................................................................................... 45

Figure 9 The welcome page and generate patient ID. ...................................................................... 46

Figure 10 Summary of the survey and summary of the first section .............................................. 47

Figure 11 Instruction page of the first section .................................................................................. 48

Figure 12 First question in the first section ..................................................................................... 49

Figure 13 Second question in the first section .................................................................................. 50

Figure 14 Third question in the first section .................................................................................... 51

Figure 15 4th question in first section .............................................................................................. 52

Figure 16 5th questions in first section ............................................................................................ 53

Figure 17 6th questions in first section ............................................................................................ 54

Figure 18 7th questions in first section ............................................................................................ 55

Figure 19 8th questions in first section ............................................................................................ 56
Figure 41 First question of this section ................................................................. 78
Figure 42 Second question of this section ............................................................. 79
Figure 43 Third question of this section ............................................................... 80
Figure 44 4th question of this section ................................................................. 81
Figure 45 5th question of this section ................................................................. 82
Figure 46 6th question of this section ................................................................. 83
Figure 47 7th question of this section ................................................................. 84
Figure 48 8th question of this section ................................................................. 85
Figure 49 9th question of this section ................................................................. 86
Figure 50 10th question of this section ............................................................... 87
Figure 51 11th question of this section ............................................................... 88
Figure 52 12th question of this section ............................................................... 89
Figure 53 Summary of fourth section ................................................................. 90
Figure 54 Introduction of fifth section ............................................................... 91
Figure 55 First question of this section ............................................................... 92
Figure 56 Second question of this section .......................................................... 93
Figure 57 Third question of this section ............................................................. 94
Figure 58 4th question of this section ............................................................... 95
Figure 59 5th question of this section ............................................................... 96
Figure 60 Summary of fifth section ................................................................. 97
Figure 61 Introduction of seventh section ......................................................... 98
Chapter 1 Introduction

Dizziness is a widespread issue, particularly among the elderly. Approximately 30% of individuals over 60 and 50% of those over 85 have reported vertigo or dizziness [1]. Data from the Hearing and Balance Clinic reveals that over 90 million Americans require medical attention for balance and dizziness problems [1,2]. Dizziness/vertigo is the second most frequently reported patient complaint, yet patients frequently disregard these symptoms [3]. The Hearing and Balance Clinic found that more than 90% of those suffering from balance and dizziness disorders never receive a diagnosis from relevant specialists such as audiologists, otologists, otolaryngologists, or neurologists [3].

Benign Paroxysmal Positional Vertigo (BPPV) is the leading cause of dizziness, but a comprehensive diagnostic method for BPPV, described in Section 1.2.1, requires significant medical resources and time. Additionally, patients may struggle to describe their BPPV symptoms accurately, leading to misdiagnoses.

This thesis addresses the challenges of BPPV diagnosis accuracy and cost by introducing a machine learning-based method utilizing a patient-facing electronic survey. The contribution of this thesis is an upgraded patients’ survey. The new survey builds upon previous work [4] by incorporating scenario-based questions, more targeted inquiries, and eliminating redundant questions. The results of the survey also reveal improved diagnostic accuracy, with a Logistic Regression Classifier achieving an accuracy of 89.8%, surpassing the previous accuracy of 74% [4]. The Decision Tree and
Naïve Bayes models also showed improved accuracy at 81.9% and 75.1% respectively, compared to previous accuracy of 70% and 72% [4].

1.1 Background

1.1.1 Benign Paroxysmal Positional Vertigo

BPPV is a peripheral vestibular disorder caused by loose crystals in the inner ear that are affected by gravity [3]. It presents as recurrent, short-term vertigo and uncontrolled eye movements (nystagmus). Although BPPV is typically self-limiting, recovery is not guaranteed, and relapses are possible. Diagnosis of BPPV requires costly medical resources [2], leading to variations in diagnostic methods across countries and differing epidemiology between nations.

The annual prevalence of BPPV is estimated at 1.6%, with a lifetime prevalence of 2.4% [5]. BPPV accounts for 20-30% of cases of vestibular vertigo, with a gender ratio of 1:2 to 1:1.5 [5]. The incidence rate increases with age, particularly after 40, as illustrated in Figure 1 [5]. The incidence of BPPV in individuals under 35 is close to zero, but rapidly increases for those over 40 and especially for those over 70.
1.1.2 Motivation

The goal of this research is to decrease diagnostic costs and complexity, prevent misdiagnosis, and improve the accuracy of the predictive model. To achieve this, an electronic patient survey has been proposed to simplify diagnosis. The results from the patient survey and a medical examination are used to train a machine learning model for BPPV prediction.

BPPV has a high rate of misdiagnosis, as seen in a study by Qian et al. where 48 out of 60 BPPV cases were misdiagnosed [1]. These misdiagnosed cases underwent standard diagnostic procedures, such as CT scans, MRI, and cerebrovascular exams,
resulting in an average misdiagnosis cost of over $1,000 [1]. Hence, this research aims to introduce methods that lower the rate of misdiagnosis.

Traditional diagnostic methods for BPPV can be complex (auditory and etiological exams), costly (CT and MRI scans), and time-consuming (static or dynamic postural tracing tests). Previous studies have shown that a patient's medical history and significant symptoms hold the key to diagnosing BPPV. Machine learning can help reduce costs by rapidly classifying potential cases, and machine learning and the electronic patient survey can minimize the time needed for accurate BPPV diagnosis.

1.1.3 Common Symptoms of BPPV

According to research conducted at Johns Hopkins, the primary symptom of BPPV is vertigo, a sensation of spinning which arises when a patient alters their head position [6]. This vertigo can cause additional problems, such as temporary loss of balance and the risk of falling when a patient moves their head, such as getting out of bed, or tilts their head forwards or backwards while walking. The duration of symptoms without medical treatment can persist for a varying amount of time, from days to weeks or indefinitely [6]. In some instances, the condition may resolve on its own.

As BPPV is an inner ear condition, ear health is also a significant symptom. Brevern et al. conducted a study with over 1000 patients and found that 10% of the patients had secondary BPPV [5]. Secondary BPPV can be caused by factors such as head or ear trauma, other ear conditions, migraines, and intubation [5]. These factors can
impact ear health, causing patients to experience inner ear pain. In the patient survey, 38 out of 50 questions pertained to symptoms, including symptom duration, ear condition, and typical symptom scenarios.

1.1.4 Diagnostic Methods for BPPV

To diagnose BPPV, the patient's medical history is gathered, and a series of physical exams are conducted, including:

- Vestibular function examinations such as Video-Nystagmography (VNG) and Rotational Chair Tests (RCT) to evaluate the patient's condition through eye movements. BPPV affects a patient's sense of balance, and VNG assesses this by monitoring eye movements.

- An audiology exam to assess hearing ability through tests with different sounds, pitches, and frequencies. This includes a pure tone audiometry test, a word recognition test, and a tympanometry test. The pure tone audiometry measures the patient's frequency range while wearing earphones, while the word recognition test evaluates their ability to understand noisy speech. The tympanometry test detects structural conditions in the ear, such as ossicle damage, inner ear tumors, and wax buildup.

- Imaging techniques, including x-rays, ultrasound, CT scans, MRI, and nuclear imaging, to visually detect dislodged otoconia, which is the root cause of BPPV.
• Balance function tests, including VNG and vestibular evoked myogenic potentials (VEMP), to evaluate a patient's sense of balance by testing muscle activity when exposed to sound from different directions.

1.2 Overview of Thesis

This thesis introduces an improved patient survey and machine learning models to optimize the cost and time required for diagnosing BPPV. Chapter 2 overviews related works and offers a comprehensive analysis of BPPV, including its pathogenesis and typical symptoms. The chapter also explains the three machine learning methods applied in this study.

Chapter 3 details the patient survey process, including its purpose and structure. Chapter 4 evaluates prior work performed by our group and highlights the unique contributions of this thesis, such as a simplified language, more precise symptom descriptions, and an in-survey tutorial to prevent misunderstandings and improve symptom reporting accuracy.

Chapter 5 validates the ideas presented in Chapter 4 through a verification process. The final chapter concludes the entire thesis, offering recommendations for future research and discussing the limitations of the current study.
Chapter 2 Background for BPPV

2.1 Prior Work

This section highlights prior research on survey-based disease classification. Auramo et al. developed an Expert System (ES) for the diagnosis of otoneurological conditions [12]. It gathers patient medical history and clinical measurement data, emulating a specialist's evaluation to determine the patient's diagnosis. For instance, migraine and BPPV both result in dizziness, but they differ in their triggering methods and symptom duration. The ES can differentiate between the two based on these distinct characteristics. Mira et al. created an ES model to classify different types of dizziness [13]. It comprises basic information about a range of dizziness-causing diseases to aid in their classification. When a patient seeks medical attention for dizziness, the doctor can use the ES model presented in [13] to make a diagnostic decision. Over 200 dizziness cases have been tested, and this ES model is widely utilized in the otoneurological field.

ONE is an expert system for diagnosing vertigo [14]. Its purpose is to provide educational support and implement a database of patient information. The ONE database contains medical history, physical symptoms, and exam results of patients who exhibit vertigo-related symptoms. Kentala et al. used pattern recognition on the ONE database, weighing and scaling vertigo-related symptoms, physical symptoms, and exam results to classify vertigo-related diseases [14]. The most probable disease is identified based on disease profiles, and fuzzy logic was employed to handle missing information and uncertainties. The accuracy of ONE was compared with that of physicians, with the mean
physician accuracy being 69% and ONE's accuracy being 65% [14].

Bansal developed Computer-Aided Diagnosis in Neurotology (CADINO), which encompasses over 100 causes of dizziness and aids clinicians in making a diagnosis [15]. CADINO is not an AI agent, but a collection of characteristics for different diseases causing dizziness. Clinicians make decisions based on these characteristics. Quick Medical Reference (QMR), a computer-assisted consultation service, also achieves similar diagnostic sensitivity as an assistant service [16]. QMR provides valuable, alternative suggestions and was rated as helpful for education [16]. [12], [13], [15]. [16] provides the foundational understanding of diagnosing vertigo, meaning that a set of symptom attributes can be used to classify it. These four expert system models demonstrate that the various causes of dizziness, differing in symptoms, can be classified by a well-trained model. A questionnaire based on the characteristics of the expert system can be used to gather training data.

Yu et al. created a questionnaire-based learning model to diagnose vertigo [17]. Over 1600 patients participated in the survey. Several models were used to classify five different types of vertigo, and nine learning algorithms were tested. The light gradient boosting machine performed best with an accuracy of 93.7% [17].

Feil et al. designed another survey-based learning model to classify six dizziness-related diseases [18]. The classification accuracy was 97.5% for Bilateral Vestibulopathy, 95.3% for Meniere's disease, 93.4% for Benign Paroxysmal Positional Vertigo, 97.5% for
Downbeat Nystagmus Syndrome, 88.3% for Vestibular Migraine, and 96.6% for Phobic Postural Vertigo [18].

Masankaran et al. introduced the Dizziness Handicap Inventory (DHI), which measures a patient's dizziness severity [19]. DHI can differentiate between different types of BPPV, which are divided into three types based on the affected ear canal: AC-BPPV (anterior canal), PC-BPPV (posterior canal), and HC-BPPV (horizontal canal). DHI can aid clinicians in determining the appropriate treatment for each type of BPPV.

Masankaran et al. investigated the performance of machine learning prediction models incorporating DHI. To differentiate between PC-BPPV and HC-BPPV, several machine learning techniques were examined, including random forests, support vector machines, k-nearest neighbors, and Naive Bayes, which were applied to a 25-question survey based on data from 114 BPPV patients. Naive Bayes performed the best with an accuracy of 73.9% [19].

Previous work on BPPV diagnosis by our group was conducted by Fasae [4] and Richburg [20]. They employed a machine learning prediction approach using patient survey data. Fasae used an ANN model with an autoencoder that transformed attributes into 22 features, with a reconstruction mean squared error of 0.11 [4]. Logistic regression, naïve Bayes, and decision tree classifiers were used to predict BPPV after correlation feature selection with similar accuracy of 70-74% [4]. Friedland et al. developed a paper survey with 162 questions focused on dizziness perception, migraine symptoms, otologic
problems, and medical test results [27]. Richburg designed three versions of the patient survey, reducing the number of questions to 84 based on Friedland's survey [27] and eliminating low-information gain questions. The third version was used to evaluate the importance of each survey question with little impact on the survey structure. Richburg's research showed that a decision tree with filters or wrappers is effective in identifying BPPV [20]. This thesis presents a new patient survey that provides data for the machine learning model, based on the research of Fasae and Richburg.

2.2 Description of BPPV

BPPV is an inner ear disorder that often leads to vertigo, a type of dizziness where one feels as though the room is spinning. The disorder is caused by dislodged otoconia (crystals) interfering with hair cells in the cupula, as shown in Figure 2, which depicts the anatomy of the inner ear. The otoconia can separate and move from the utricle to the posterior semicircular canal and, in rare cases, to the anterior or horizontal semicircular canal. When a person moves their head, the displaced otoconia activates the hair cells in the cupula, which sense body direction. The activated otoconia may then send incorrect signals to the brain, causing an erroneous perception of balance and resulting in dizziness.
The dislodged otoconia in the cupula can shift when the patient moves their head and touches the hair cells. As a result, a BPPV episode is often triggered by a change in gravity affecting the patient's head, leading to a sudden and temporary experience of dizziness. Symptoms frequently include nausea, vertigo, lightheadedness, a feeling of floating, loss of balance, and visual hallucinations [21].

2.3 Machine Learning Algorithms

Machine learning algorithms can be classified into three groups: supervised, unsupervised, and reinforcement. Supervised learning involves examples with accompanying labels, while unsupervised learning lacks labels. The objective in supervised learning is to establish a mapping between examples and labels. For
unsupervised learning, the goal is to find a compact representation of the data.

Reinforcement learning is similar to supervised learning, but labels are only provided at certain intervals and not for every example. This study employs supervised learning, as each patient's features have labels indicating a diagnosis of BPPV or not. This thesis employs three standard supervised learning algorithms (logistic regression, naïve Bayes, and decision tree) for the BPPV diagnosis problem, which are described below.

2.3.1 Candidate Algorithms

This section outlines the selection of naïve Bayes, Decision Tree, and Logistic Regression algorithms for the study. The survey data collected from BPPV patients consisted of 50 questions and demographic information such as age and gender, resulting in 52 input factors for the machine learning model. The goal was to predict the presence of BPPV. The Decision Tree classifier was preferred over others due to its ease of implementation and data preparation, as all survey questions could be used as tree nodes without the need for normalization or scaling. Furthermore, the visual representation of the tree makes it easy to understand.

On the other hand, the naïve Bayes classifier requires less training data compared to others. In this case, it calculates the posterior probability of a patient having BPPV based on the probability of patients with BPPV, the probability of patients with the same factors (50 survey questions, age and gender), and the probability of patients with the same factors among all BPPV patients. This reduces the amount of data required for
accurate calculation.

Lastly, Logistic Regression was chosen for its simplicity in training and implementation. The model uses each factor and its coefficient to calculate the output and evolves the coefficients during training to arrive at the optimal solution. Additionally, the model displays the positive or negative impact of each factor on the output, making it intuitive to understand.

2.3.2 Logistic Regression Classifier

Logistic regression is a classification method that is used for binary, probabilistic classification [23]. For example, the probability that a person has BPPV based on their answers to a survey. The result is often used for weighted summation with other eigenvalues instead of direct multiplication.

We examine the logistic function, also known as sigmoid function and logistic distribution. The sigmoid function is a continuous probability distribution defined as

\[ y = \frac{1}{1 + e^{-z}} \quad (1) \]

Figure 3 shows the logistic distribution.
As shown in Figure 3, when variable $z$ is greater than 0, $y$ is greater than 0.5. We assume $y$ represents probability. When the probability of $y$ is greater than 0.5, we assign input $z$ as label “1”. When probability $y$ less than 0.5, we assign input $z$ as label "0". In binary classification, inputs are a series of parameters that can be written as:

$$z = \omega_0 + \omega_1 x_1 + \omega_2 x_2 + \cdots + \omega_n x_n,$$

(2)

where $x_0, x_1, \cdots, x_n$ are attributes for $z$ and $\omega_0, \omega_1, \omega_2, \cdots, \omega_n$ are weights for the attributes.

When training the model, a cost function is needed for labels “1” and “0”. When $label = 0$, the cost function is

$$cost = -\ln(1 - h).$$

(3)
When \( label = 1 \), the cost function is

\[
\text{cost} = -\ln(h). \tag{4}
\]

For both cases (\( label = 1 \) or 0), the cost function written as,

\[
\text{cost} = -y^* \ln(h) - (1 - y^*) \ln(1 - h), \tag{5}
\]

where \( y^* = label \).

The training process estimates \( \omega_0, \omega_1, \omega_2, \ldots, \omega_n \) to minimize the cost function.

The pseudocode for the Logistic Regression Classifier is shown below.

---

**Algorithm-1: Logistic Regression**

**Input:** Training Data

1. Label vector \( label = [0,1] \)

2. For \( i \) from 1 to \( n \)

3. Calculate: \( z = \omega_0 + \omega_1 x_1 + \omega_2 x_2 + \cdots + \omega_n x_n \)

4. Based on the value \( z \) calculated from step 1 calculate:

5. \[ y = \frac{1}{1+e^{-z}} \]

6. Minimize cost function

7. \[ \text{cost} = -y^* \ln(h) - (1 - y^*) \ln(1 - h) \]

8. if \( P(1|d_j) > 0.5 \),

9. \( label = 1 \)

10. otherwise assign label 0.
2.3.3 Naïve Bayes Classifier

The Naïve Bayes Classifier uses probability statistics to classify samples. The combination of prior probability and posterior probability in a naïve Bayes classifier avoids the subjective bias of using only prior probability [24]. Because of a solid mathematical foundation, the misjudgment of naïve Bayes Classifier is low [24]. The Bayesian classification algorithm shows high accuracy when the data set is large, and the algorithm itself is straightforward [25].

The Bayes function calculates the posterior probability. The Bayes function solves the problem often encountered in real life. The following is the Bayes Formula:

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}.$$  \hspace{1cm} (6)

The posterior probability of A can be calculated by the probability of A, the posterior probability of B, and the probability of B. We consider B as a factor of A. When A determined by a series of factors $B_1, B_2, B_3, ..., B_n$, the posterior probability of A can be calculated as

$$P(A|B_1, B_2, B_3, ..., B_n) = \frac{P(A)\prod_{i=1}^{n} P(B_i|A)}{\prod_{i=1}^{n} P(B_i)},$$  \hspace{1cm} (7)

which is determined by probability of A, probability of B, and posterior probability of B. In BPPV, for example, assuming symptoms and triggering methods are two factors of whether a patient has BPPV, we assign them as $B_1$ and $B_2$ and A is “has BPPV”. Then, the posterior probability becomes $P(\text{has BPPV}|B_1, B_2)$. Based on formula (7), the
components for calculating the probability of a patient having BPPV are the probability of all patients having BPPV, the probability of BPPV patients who have the same condition $B_1$ and $B_2$, and probability of patients who have the same conditions $B_1$ and $B_2$. These three components can be calculated from the dataset.

The following is the pseudocode of Naïve Bayes Classification.

Algorithm-2: Naïve Bayes Classification

Input: Training dataset $T$,

$$F = (f_1, f_2, f_3, f_4, \ldots, f_n)$$

// value of the predictor variable in the testing dataset.

Output:

A class of testing dataset.

Steps:

1. Read the training dataset $T$.

2. Calculate the mean and standard deviation of the predictor variables in each class.

3. Repeat

4. Calculate the probability of $f_i$ using the gauss density equation in each class.

5. Until the probability of all predictor variables ($f_1, f_2, f_3, f_4, \ldots, f_n$) has been calculated.

6. Calculate the likelihood for each class.

7. Get the greatest likelihood.
2.3.4 Decision Tree Classifier

A decision tree is an algorithm used for classification and regression. It has a tree-shaped structure. As shown in Figure 4, starting from the root node, the decision tree tests a feature sample and assigns the instance to its child nodes according to the test results. Each child node corresponds to a value of the feature, so the samples are tested and assigned recursively until the leaf node is reached. Finally, the sample is assigned the leaf node's class [26].

![Figure 4 Decision Tree Structure](image)

In decision tree classification, "find best split" refers to the process of identifying
the attribute and its corresponding value that produces the highest information gain (or the lowest impurity) when we use it to split the data at a particular node.

The information gain is a measure of the reduction in entropy (or impurity) achieved by splitting the data on a particular attribute. The attribute with the highest information gain is considered the best split because it provides the most discriminatory power to separate the different classes in the dataset.

The decision tree algorithm iterates over all attributes, and their possible values to determine that split that produces the highest information gain. Once the best split is identified, the data is partitioned into two or more subsets based on the split, and the process is repeated recursively on each subset until a stopping criterion is met, such as a maximum tree depth or a minimum number of samples at each leaf node. The quality of the split can have a significant impact on the performance of the decision tree classifier, and finding the best split is a critical step in building an accurate and efficient decision tree model.

The pseudocode of a decision tree classifier is shown below.

---

**Algorithm-3: Generate a Decision Tree**

1. GenDecTree (Sample S, Features F)

2. If stopping_condition(S, F) = true then

3. Leaf = createNode()

4. leafLabel = classify(s)

5. return leaf
6. root = createNode()

7. root.test_condition = findBestSplit(S,F)

8. V = \{v \mid v \text{ a possible outcome of root.test_condition}\}

9. For each value \( v \in V \):

10. \( S_v = \{ s \mid \text{root.test}_\text{condition} = v \text{ and } s \in S \} \);

11. Child = TreeGrowth \( (S_v, F) \)

12. add child as descent of root and label the edge \( \{\text{root} \rightarrow \text{child}\} \) as \( v \)

13. return root

2.3.5 Cross-validation

Cross-validation is a method used to evaluate the performance of a machine learning model by using a subset of the available data to train the model and a different subset to test the model. The main idea behind cross-validation is to use multiple rounds of training and testing to get a more accurate estimate of the model's performance.

K-fold cross-validation is the most popular cross-validation method. In k-fold cross-validation, based on the amount of data, the data is divided equally into k folds. Then, the learning model is trained and tested k times. In every round, a different fold is used as the test set. The remainder of the k-1 folds are the training set. The performance (such as accuracy or error rate) of each round is recorded.

After k rounds of training and testing, the performance metrics from each round are
averaged to obtain a more robust estimate of the model's performance. This approach is beneficial because it allows the use of all available data for training and testing and helps to avoid overfitting by testing the model on data that it has not seen during training.

Other forms of cross-validation include leave-one-out cross-validation, which involves training the model on all but one of the data points and testing it on the remaining point, and stratified cross-validation, which ensures that the distribution of classes is preserved in each fold. Overall, cross-validation is a powerful tool for evaluating the performance of a machine learning model and is widely used in practice to select and tune models.
Chapter 3 Survey Process

3.1 Aim

The patient survey used in this study was designed based on the symptoms and characteristics of BPPV as described in section 2.2. The questions in the survey are based on previous studies [20], [27]. Friedland et al. developed a 162-question paper-based questionnaire [27], which focused on dizziness perception, migraine symptoms, otologic problems, and medical test results. Richburg et al. [20] then reduced the number of questions to 84 and further to 41, removing those deemed unimportant through analysis (by calculating the information gain from each question). The survey was further updated with questions based on features collected by Siermala et al. [28]. The third version of Richburg's survey was used to evaluate the importance of each question, with little change made to the survey structure.

The details of the updates made to the survey questions in this thesis are discussed in Chapter 4. The survey is implemented as an Android-based app and includes questions related to symptoms and disease characteristics. All questions of the new survey are shown in Appendix A. The answers provided by patients in the survey serve as the data used to train the machine learning models.

The survey generates a database of answers to all questions, and the goal of the machine learning model is to match the clinician's diagnosis.
3.2 Structure

The patient survey is composed of seven parts, which are the perceptual description of the symptoms, the self-judgment of the symptoms (the patient's self-diagnosis according to the medical definitions), the duration of the symptoms, the triggering methods, ear-related problems, migraine, and related medical history. There are 50 questions, and at the end of each section there is a summary paragraph to confirm the patient’s answers.

3.2.1 Questions

The first part (ten questions) investigates the patient's symptoms. These ten questions were summarized by Dr. Friedland from the Medical College of Wisconsin through the patient's self-description when he treated BPPV patients. Most scenarios of questions designed mainly summarized based on the patients’ self-description. The questions are general and colloquial. The purpose of this section is to eliminate the lack of medical knowledge that may cause patients to misunderstand the questions.

The second part (five questions) asks about four specific symptoms of vertigo. Patients choose their symptoms through medical definitions (the medical definitions are shown at the beginning of this section). The medical definition may cause misunderstanding of symptoms.

The third part (six questions) investigates the duration of the patient’s symptoms. The duration is divided into a few minutes, 20 minutes to an hour, within twelve hours,
within one day, and lasts longer than one day. This part helps to determine the duration of the patient's symptoms.

The fourth part (12 questions) investigates the triggering methods of the patient's symptoms. Since BPPV is an ear-related disease, moving the head in different directions, strengths, and frequencies will trigger vertigo [3]. Moreover, different movements can also reflect the severity of BPPV. Therefore, the 12 questions in the fourth part not only ask whether different actions will cause vertigo, but also ask whether the action will make the symptoms worse.

The fifth part (five questions) asks about the specific conditions of the patient’s ears. The main questions are whether the patient feels pain in the ears, whether they hear specific sounds when vertigo occurs, and whether the two ears would have similar symptoms.

The sixth part (10 questions) investigates the migraine-related symptoms. These include the problem of migraine symptoms, whether a migraine has been diagnosed before, and whether dizziness occurs under certain circumstances. Since the symptoms of migraine are like BPPV, this part is used to distinguish these two diseases.

The seventh part (two questions) investigates the patient's medical history (chemotherapy and hip or knee replacement surgery)

3.2.2 Scenario

The patient presses the “Start” button to enter the survey and generate a user's ID
number by the system. The user’s ID number is used in the database to sort and identify each patients’ answer. The welcome interface tells the user the total number of questions and reminds the user to answer carefully. In the first part, the patient chooses very similar, similar, unrelated, dissimilar, and very dissimilar through 10 questions about different symptom descriptions. At the beginning of the second part, the user is provided four different medical definitions of vertigo, and in the follow-up questions, they select the one that best describes their symptoms. Then, they complete all the questions and answer them in turn. Before the beginning of each part, the user will be informed of the main content and quantity of the questions. At the end of each part, a summary page is generated to help the user to confirm that there is no incorrect answer in the current section. After going over all questions, the system will automatically record the answers to all questions, generate a database, and send it to the researchers by email.

3.3 Improvements of Patient Survey

This chapter presents enhancements made to the patient survey. The goal is to gather more precise information and eliminate redundant or unhelpful questions. The improvements include a clearer description of certain questions, and improvement in the wording of the survey. The following subsections outline the main methods used to improve the survey, including simplification of the survey structure, improvement in question wording, addition of an in-survey tutorial, and mitigation of inaccurate symptom reporting.
3.3.1 Survey Simplification

To improve the patient survey, simplification of language and elimination of redundant questions was necessary. Complex and lengthy descriptions could impact the user's accuracy in judging symptoms and recalling medical history. To simplify the survey, language was made simpler, sections were reordered, and duplicate questions were removed. Language simplification involved reducing lengthy explanations, reordering sentences to simplify them, and eliminating duplicated explanations. For instance, in the explanation of the four medical terms related to vertigo, only the definitions were kept, while examples were omitted to avoid misunderstandings.

In the fifth section (conditions of the patient’s ears), one question was deleted. This question asked that “Do you have pain in both ears equally?”. Because the main reason of BPPV is dislodged otoconia, equal conditions happen to both ears are rarely appearing.

3.3.2 In-survey Tutorial

Benign paroxysmal positional vertigo (BPPV) is a common condition that affects the inner ear, resulting in brief episodes of dizziness or vertigo. Patients experiencing BPPV often visit doctors to seek medical attention. However, patients may not clearly understand their symptoms, which can lead to misunderstandings about their condition.

This lack of understanding can lead to inaccurate descriptions of symptoms and misinterpretation of medical definitions. For example, patients may use different words or
phrases to describe their symptoms, making it difficult for doctors to accurately diagnose and treat the condition. In addition, patients may not understand medical terms used by doctors, which can lead to confusion and further misunderstanding.

To address these issues, the survey includes sections designed to provide necessary explanations and guidance. These sections may include information about BPPV and its symptoms, as well as descriptions of medical terms and definitions used to diagnose and treat the condition. By providing this information, patients can better understand their symptoms and communicate them effectively to their doctors. This can lead to more accurate diagnoses and better treatment outcomes for patients with BPPV.

In Section 2, four types of vertigo are introduced with a dedicated page for patients to learn the definitions, promoting a better understanding of medical terms. Additionally, Section 3, which inquiries about symptom duration, presents two distinct scenarios to clarify duration measurement for the patient. For instance, if a patient experiences dizziness for a minute after turning over in bed and again a few minutes later, the duration of their symptoms would be considered a minute, not a few minutes.

In Section 3, the duration of patients’ symptoms was asked. The duration of their symptoms from a few seconds to one or two days. Six questions in this section ask different interval duration, which are “seconds to a few minutes”, “between 20 minutes and an hour”, “more than an hour but less than 12 hours”, “more than 12 hours but not more than 1 day”, and “more than a day”. If a patient’s duration of symptoms lasts
several hours, the patient should choose “Yes” on the question “Does your symptoms last more than an hour but less than 12 hours”. However, in this case, the duration of a couple of hours includes “a few minutes” and “20 minutes to an hour” and question ask about “a few minutes” comes before question ask that “Does your symptoms last more than an hour but less than 12 hours”. Thus, patients may select the wrong answer. Therefore, two tutorial pages show this possible scenario at the beginning of this section.

3.3.3 Descriptions of Symptoms

In some cases, patients may still have difficulty accurately describing their symptoms despite the in-survey tutorial. When this happens, a summary of common symptoms reported by other patients with similar conditions can be used as a substitute for a doctor's explanation. This approach can help patients better understand their symptoms and provide doctors with valuable information to aid in diagnosis and treatment.

The new patient survey includes a section specifically designed to address this issue. This section was compiled by Dr. Friedland of the Medical College of Wisconsin and summarizes common symptoms reported by patients with BPPV. The section consists of 10 questions, each based on a patient's statement about their symptoms. Patients will select the appropriate level of similarity (very similar, similar, unrelated, dissimilar, or very dissimilar) for each question.

For example, one question may ask about the duration of a patient's symptoms,
with options ranging from a few seconds to several minutes. Another question may ask about the severity of the symptoms, with options ranging from mild to severe. By answering these questions, patients can better understand their symptoms and how they compare to those reported by other patients with BPPV.

This approach can be particularly useful for patients who may have difficulty communicating their symptoms to their doctor due to language barriers or other factors. By providing a summary of common symptoms, patients can still receive valuable information about their condition and potential treatment options. Ultimately, this can lead to better outcomes for patients with BPPV and other conditions.

3.4 Survey Questions

This section lists all the survey questions used in the patients’ survey. The first survey section has 10 questions which mainly summarized based on the patients’ self-description. Patients should scale their similarity to the scenarios in each question which are “very similar”, “similar”, “neither dissimilar or similar”, “dissimilar”, and “very dissimilar”. Then, except for the first question in the second section, choose a most likely symptom, all the questions in the rest part are choosing “yes” or “on”.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Question Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section1 q1</td>
<td>My dizziness often comes when lying down or turning in bed. It</td>
</tr>
<tr>
<td>Question Number</td>
<td>Question Text</td>
</tr>
<tr>
<td>-----------------</td>
<td>--------------</td>
</tr>
<tr>
<td></td>
<td>feels like I am spinning or moving. The dizziness lasts less than a minute. This has happened on many days or weeks.</td>
</tr>
<tr>
<td>Section1 q2</td>
<td>My dizziness comes at any time. I could be sitting still, and it happens. I feel lightheaded and woozy. It lasts seconds to minutes.</td>
</tr>
<tr>
<td>Section1 q3</td>
<td>My dizziness often comes when moving my head such as lying down in bed, tilting my head back, or reaching and bending. It is a strong dizzy feeling. The strong feeling lasts less than a minute, but I do not feel well or normal for longer.</td>
</tr>
<tr>
<td>Section1 q4</td>
<td>My dizziness comes when I sit or stand up. I feel lightheaded like I might pass out. It lasts anywhere from a few seconds to a few minutes.</td>
</tr>
<tr>
<td>Section1 q5</td>
<td>My dizziness usually occurs when I am walking. It is made worse if I turn my head left and right. It feels like I drift from side to side.</td>
</tr>
<tr>
<td>Section1 q6</td>
<td>My dizziness is always there. I feel better when lying in bed. I feel worse when doing things like walking down aisles in a store or going down long hallways.</td>
</tr>
<tr>
<td>Section1 q7</td>
<td>My dizziness feels like I am unsteady. It is worse in the dark. I feel like I drift to the side when walking.</td>
</tr>
<tr>
<td>Section1 q8</td>
<td>My dizziness comes on suddenly without warning. It lasts hours but not more than a day. The dizziness makes me feel like throwing up. When I get the dizziness, I cannot do anything except sit still or lie down.</td>
</tr>
<tr>
<td>Section1 q9</td>
<td>My dizziness comes at any time. It can last hours or even days. I often get a headache when I am dizzy. I feel very tired and often lie down in the dark.</td>
</tr>
<tr>
<td>Section1 q10</td>
<td>I get short episodes of dizziness. The dizziness comes when I move. This happens many times in a day or week.</td>
</tr>
<tr>
<td>Section2 q1</td>
<td>Choose type of dizziness that best characterizes your condition. Using the definitions noted on the prior page, which one of the following most closely resembles your main form of dizziness.</td>
</tr>
<tr>
<td>Section2 q2</td>
<td>Have your symptoms occurred once or more than once?</td>
</tr>
<tr>
<td>Section2 q3</td>
<td>With your symptoms have you had nausea or felt queasy?</td>
</tr>
<tr>
<td>Section2 q4</td>
<td>During your symptoms have you ever had blurry vision?</td>
</tr>
<tr>
<td>Section2 q5</td>
<td>During your symptoms have you ever had double vision?</td>
</tr>
<tr>
<td>Section3 q1</td>
<td>Do your symptoms last between seconds to a few minutes?</td>
</tr>
<tr>
<td>Section3 q2</td>
<td>Do your symptoms last between 20 minutes and an hour?</td>
</tr>
<tr>
<td>Section3 q3</td>
<td>Do your symptoms last more than an hour but less than 12 hours?</td>
</tr>
</tbody>
</table>
| Section3 q4     | Do your symptoms last more than 12 hours, but not more than 1
<table>
<thead>
<tr>
<th>Question Number</th>
<th>Question Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section3 q5</td>
<td>Do your symptoms last more than a day?</td>
</tr>
<tr>
<td>Section3 q6</td>
<td>Is your symptom currently with you 24 hours a day, never stopping?</td>
</tr>
<tr>
<td>Section4 q1</td>
<td>Is your symptom typically made worse or triggered by sitting up or standing up?</td>
</tr>
<tr>
<td>Section4 q2</td>
<td>Is your symptom typically made worse or triggered by reaching or bending?</td>
</tr>
<tr>
<td>Section4 q3</td>
<td>Is your symptom typically made worse or triggered by tilting your head back?</td>
</tr>
<tr>
<td>Section4 q4</td>
<td>Is your symptom typically made worse or triggered by lying down in bed?</td>
</tr>
<tr>
<td>Section4 q5</td>
<td>Is your symptom typically made worse or triggered by rolling over or turning in bed?</td>
</tr>
<tr>
<td>Section4 q6</td>
<td>Is your symptom typically made worse or triggered by turning your head while walking?</td>
</tr>
<tr>
<td>Section4 q7</td>
<td>Is your symptom typically made worse or triggered by supermarket aisles?</td>
</tr>
<tr>
<td>Section4 q8</td>
<td>Is your symptom typically made worse or triggered by seeing things go past over and over like scrolling on your phone or watching a train pass?</td>
</tr>
<tr>
<td>Section4 q9</td>
<td>Is your symptom typically made worse or triggered by loud noises?</td>
</tr>
<tr>
<td>Section4 q10</td>
<td>Is your symptom typically made worse or triggered by automobile rides?</td>
</tr>
<tr>
<td>Section4 q11</td>
<td>Is your symptom typically made worse or triggered by walking on uneven ground?</td>
</tr>
<tr>
<td>Section4 q12</td>
<td>Is your symptom typically made worse or triggered by driving a car at night?</td>
</tr>
<tr>
<td>Section5 q1</td>
<td>Do you have ear pain?</td>
</tr>
<tr>
<td>Section5 q2</td>
<td>Do you get an ear infection at least once a year?</td>
</tr>
<tr>
<td>Section5 q3</td>
<td>Do you have pain in ONE EAR ONLY or ONE EAR MORE THAN THE OTHER?</td>
</tr>
<tr>
<td>Section5 q4</td>
<td>Do you hear a ringing, buzzing, hissing, or other noise in ONE EAR ONLY or ONE EAR MORE THAN THE OTHER?</td>
</tr>
<tr>
<td>Section5 q5</td>
<td>Do you hear a ringing, buzzing, hissing, or other noise in BOTH EARS EQUALLY (or in your head)?</td>
</tr>
<tr>
<td>Section6 q1</td>
<td>Have you had a total of 5 or more bad headaches in your lifetime?</td>
</tr>
<tr>
<td>Section6 q2</td>
<td>Have you ever had a headache that throbs or pulses?</td>
</tr>
<tr>
<td>Question Number</td>
<td>Question Text</td>
</tr>
<tr>
<td>-----------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Section6 q3</td>
<td>Have you ever had increased sensitivity to light with a headache?</td>
</tr>
<tr>
<td>Section6 q4</td>
<td>Have you ever had increased sensitivity to sounds with a headache?</td>
</tr>
<tr>
<td>Section6 q5</td>
<td>Have you ever had increased sensitivity to smells with a headache?</td>
</tr>
<tr>
<td>Section6 q6</td>
<td>Have you ever had nausea or vomiting with a headache?</td>
</tr>
<tr>
<td>Section6 q7</td>
<td>Have you ever had your symptoms associated with a headache?</td>
</tr>
<tr>
<td>Section6 q8</td>
<td>Have you ever been diagnosed with migraine headaches?</td>
</tr>
<tr>
<td>Section6 q9</td>
<td>Have you ever been diagnosed with ocular or retinal migraines (means a temporary change in vision that may come with a headache)?</td>
</tr>
<tr>
<td>Section6 q10</td>
<td>Did you get carsick as a child?</td>
</tr>
<tr>
<td>Section7 q1</td>
<td>Have you had a hip or knee replacement?</td>
</tr>
<tr>
<td>Section7 q2</td>
<td>Have you had chemotherapy?</td>
</tr>
</tbody>
</table>
Chapter 4 Results

4.1 Data Analysis

This section introduces the results of the machine learning prediction models. The data used for training the machine learning model are collected from 40 patients who visited Medical College of Wisconsin from April to September in 2021 who were suspected of having BPPV. All patients were offered the opportunity to take the survey. A database, including all survey answers, was generated. The label (whether they have BPPV) was determined from a medical examination. Data were collected from human subjects under approved institutional review board agreements. The data stored in the results database was anonymized.

All 40 patients' age distribution is shown in Figure 4 (orange line is the positive rate of BPPV). The interval of age from 67 to 81 has the biggest population, which is 18 cases and 45% of all patients.
Figure 5 Age distribution of 40 patients

Figure 6 shows the gender situation of all 40 patients (21 male and 19 females).

Figure 7 shows the positive rate of male and female. Female has a 63% positive rate which is higher than male’s positive rate of 52%.
4.2 Machine Learning Results

The prediction model uses three algorithms, which include decision tree, naive Bayes, and logistic regression. The results after cross-validation can significantly improve accuracy. Cross-validation proposes one sample in turn among 40 samples and processed by machine learning algorithms. Finally, the best combination of the remaining 11 samples is obtained which has the highest accuracy.

Table 2 shows the results of three algorithms after cross validation.

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Average Training Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision Tree</td>
<td>81.9%</td>
</tr>
<tr>
<td>Logistics Regression</td>
<td>89.8%</td>
</tr>
<tr>
<td>Naïve Bayes</td>
<td>75.1%</td>
</tr>
</tbody>
</table>
Given that there are only 40 data points, leave one out cross validation was used.

The Naive Bayes Classifier performs poorly due to the need for a large dataset to calculate the posterior probability of 52 factors among all patients with BPPV. With a small dataset, the probability is not accurate. The Decision Tree requires a large dataset to obtain the highest information gain to construct an optimal tree, resulting in limited classification ability when the dataset is small. Logistic Regression, however, does not require a large dataset to calculate the posterior probability. With 40 data points, the coefficients evolve 40 times, reflecting the significance of each factor, leading to its superior performance among the three models.

4.3 Discussion

The dataset comprises 40 patients' survey answers. Upon analysis of 50 questions from the 40 patients, it was determined that the machine learning model may be overfitting. Using part of the patient information as a test set results in limited data available for model training, leading to potentially inaccurate results.

The patient background information in the survey primarily encompasses medical history, age, and gender. For BPPV, height, weight, and preferred sports are also relevant information. As discussed in Section 2.2, changes in head position can induce symptoms related to BPPV, such as vertigo. For instance, lifting weights results in vertical head movement, while playing basketball involves intense body movement that leads to head movement in various directions, mostly horizontal. Thus, it is advisable to gather more
patient background information.

In machine learning, factors such as age and gender are not deemed as significant characteristics due to their limited impact among the 50 other questions. The small dataset of 12 individuals, representing diverse ages and genders, may not reflect the independence of age and gender, contributing to low accuracy. However, dividing the sample into two parts by gender is not likely to improve accuracy significantly.
Chapter 5 Conclusion and Future Work

This thesis presents a machine learning approach for BPPV classification through a patient survey data collection. A key contribution of this study is the improvement of the patient survey questionnaire and addressing its limitations, as well as offering suggestions for future BPPV prediction models. The revised survey, through the elimination of redundant questions and the addition of a section on patients' symptoms, resulted in a significant improvement in classification accuracy compared to prior research (70-74%) [4]. Among the three machine learning algorithms tested, logistic regression achieved the highest accuracy rate of 89.8%.

The recent enhancements to the survey, which included elimination of redundant questions and inclusion of a new section on patient symptoms, resulted in a significant improvement in accuracy. However, there is a large variability across cross-validation folds. This can be attributed to a small sample size, the absence of background questions, and high inter-question correlation, which will be discussed in more detail in the next sections.

Next, we suggest future work that may improve the classification results. The patient survey encompasses 50 questions and considers the patient's age and gender as two of the 52 factors, along with the results of medical tests, to be used as features in the machine learning model. However, not all 52 factors hold equal importance. Questions from the first and second sections of the survey ask about the patient's symptoms, with
the first section based on the patient's own description and the second section based on
the medical definition of symptoms. Self-reported symptoms may not be as significant as
symptoms assessed with medical definitions. For instance, in the machine learning
process, age is treated as just another column alongside the results of the fifty questions.
But in medical literature, individuals over 45 years old are more likely to develop BPPV
[11]. Increasing the weight of age during the machine learning model training may
enhance prediction accuracy.

It is recommended to include patients without dizziness symptoms in the patient
population. Currently, the data collected pertains only to patients with ear or head related
issues. Including healthy patients in the sample would provide a better characterization of
the healthy population by the machine learning model and provide more data for
researchers.

As previously noted in section 4.3 the limited patient background information of
only medical history, age, and gender may not be sufficient. Including additional factors
such as height, weight, and preferred sports activities could enhance the accuracy of the
model.
Bibliography


https://epublications.marquette.edu/electric_fac/514


https://doi.org/10.1001/jamaoto.2015.3663

"Evaluation and classification of otoneurological data with new data analysis methods 
Appendix A

The appendix shows screenshots of every page in the patient’s survey. The first step is creating a new database for recording new patient’s answer and generating a Patent Record ID for the patient in the database (Figure 9). Figure 10 is the instruction page of the survey. It tells the number of questions.

Figure 8 shows the home page of the survey.
It appears you need to send records. Press 'Send Records'.
Then press 'Create New Record' to create a new patient record.
Figure 9 shows the welcome page of the survey, and a patient record ID is generated at the back.

![Welcome Page](image)

Figure 9 The welcome page and generate patient ID.
Figure 10 is the instruction page of the survey. It tells the number of questions.

Welcome to the survey. We appreciate your participation. The survey has six sections and a maximum of 50 questions.

Please answer all questions before returning the tablet to the study coordinator.
Figure 11 is the instruction page of the first section. The number of questions is shown. Figure 8-12 are the questions in the first section, which provide scenarios of common symptoms for patients to choose.

![Screen capture of a survey interface](image)

Figure 11 Instruction page of the first section
Figure 12 shows the first question in the first section.

My dizziness often comes when lying down or turning in bed. It feels like I am spinning or moving. The dizziness lasts less than a minute. This has happened on many days or weeks.

- Very Similar
- Similar
- Neither dissimilar or similar
- Dissimilar
- Very dissimilar

Figure 12 First question in the first section
Figure 13 is the second question in the first section.

My dizziness comes at any time. I could be sitting still, and it happens. I feel lightheaded and woozy. It lasts seconds to minutes.

○ Very Similar
○ Similar
○ Neither dissimilar or similar
○ Dissimilar
○ Very dissimilar

Figure 13 Second question in the first section
Figure 14 is the third question in the first section.

My dizziness often comes when moving my head such as lying down in bed, tilting my head back, or reaching and bending. It is a strong dizzy feeling. The strong feeling lasts less than a minute, but I do not feel well or normal for longer.

- Very Similar
- Similar
- Neither dissimilar or similar
- Dissimilar
- Very dissimilar
Figure 15 is the 4th question in the first section.

My dizziness comes when I sit or stand up. I feel lightheaded like I might pass out. It lasts anywhere from a few seconds to a few minutes.

- Very Similar
- Similar
- Neither dissimilar or similar
- Dissimilar
- Very dissimilar
Figure 16 is the 5th question in the first section.

My dizziness usually occurs when I am walking. It is made worse if I turn my head left and right. It feels like I drift from side to side.

- Very Similar
- Similar
- Neither dissimilar or similar
- Dissimilar
- Very dissimilar
Figure 17 is the 6th question in the first section.

My dizziness is always there. I feel better when lying in bed. I feel worse when doing things like walking down aisles in a store or going down long hallways.

- Very Similar
- Similar
- Neither dissimilar or similar
- Dissimilar
- Very dissimilar
Figure 18 is the 7th question in the first section.

My dizziness feels like I am unsteady. It is worse in the dark. I feel like I drift to the side when walking.

- Very Similar
- Similar
- Neither dissimilar or similar
- Dissimilar
- Very dissimilar
Figure 19 is the 8th question in the first section.

My dizziness comes on suddenly without warning. It lasts hours but not more than a day. The dizziness makes me feel like throwing up. When I get the dizziness, I can not do anything except sit still or lie down.

- Very Similar
- Similar
- Neither dissimilar or similar
- Dissimilar
- Very dissimilar
Figure 20 is the 9th question in the first section.

My dizziness comes at any time. It can last hours or even days. I often get a headache when I am dizzy. I feel very tired and often lie down in the dark.

- Very Similar
- Similar
- Neither dissimilar or similar
- Dissimilar
- Very dissimilar
Figure 21 is the 10th question in the first section.

I get short episodes of dizziness. The dizziness comes when I move. This happens many times in a day or week.

- Very Similar
- Similar
- Neither dissimilar or similar
- Dissimilar
- Very dissimilar

Figure 21 10th question in first section
Figure 22 is the instruction page of the second section. As section 3.2.1 introduced, the section provides four medical definitions of different dizziness are shown in Figure 23. Figure 24-29 are the questions in this section. A summary of this section is generated for patients to review their answers, which shows in Figure 18.

This section asks about your type of dizziness. Unless otherwise specified, please answer these questions in regards to your dizziness.

There are 5 questions in this section.
Four medical definitions of different dizziness are shown in Figure 23.

- **Vertigo** is a sensation of feeling like you are spinning or that the world around you is spinning.

- **Faint or light-headedness** is a feeling that you are about to pass out.

- **Wooziness** is a feeling of being off, like you are drunk or not stable. Some people feel foggy like they cannot think clearly.

- **Imbalance** is a sensation of feeling unsteady on your feet, you may have difficulty keeping your balance.
Choose type of dizziness that best characterizes your condition. Using the definitions noted on the prior page, which one of the following most closely resembles your main form of dizziness.

- Vertigo
- Faint or Light-Headness
- Wooziness
- Imbalance
- None of these

Figure 24 First question in second section
Figure 25 is the second question in the second section.

Has your vertigo occurred once or more than once?

- Once
- More than once

Figure 25 Second question in second section
Figure 26 is the third question in the second section.

With your vertigo, have you had nausea or felt queasy?

- Yes
- No
Figure 27 is the 4th question in the second section.

During your vertigo, have you ever had blurry vision?

- Yes
- No
Figure 28 is the last question in the second section.

During your vertigo, have you ever had double vision? Double vision is when you see a double image where there should only be one.

- Yes
- No

Figure 28 5th question in second section
Figure 29 shows the summary page of the second section.

Figure 29 Summary of second section

The most noticeable part of your symptom is vertigo. Your vertigo has occurred more than once. With your vertigo you have had nausea or felt queasy. During your vertigo you have had blurry vision and double vision.

Is this correct? If not, repeat section.

SAVE AND CONTINUE

NO. MODIFY.
The topic of the third section is duration and frequency of patients’ symptoms, which is shown in Figure 30. Two example scenarios to eliminate misunderstandings shown in Figure 31 and 32. Figure 33-38 show questions in the third section ask about patients’ duration and frequency. Figure 39 summarizes the answers in this section for patients to review.

Figure 30 Introduction of third section
Scenario 1

Example of how to answer questions based on the above scenario.

You turn over in bed and notice extreme dizziness that lasts for about 30 seconds. An hour later you turn in bed again and have the same thing. This represents two spells/episodes of dizziness lasting 30 seconds (NOT dizziness lasting an hour)

Did your dizziness last between seconds to a few minutes?

- Yes
- No

Did your dizziness last between 20 minutes and an hour?

- Yes
- No

Figure 31 Scenario 1 of how to answer questions.
Scenario 2

Example of how to answer questions based on the above scenario.

Almost every morning when you get out of bed you have severe dizziness lasting about a minute. This goes on every day for two weeks. This represents spells/episodes of dizziness lasting about a minute (NOT dizziness lasting 2 weeks)

Did your dizziness last between seconds to 1 minute?
- Yes
- No

Did your dizziness last more than a day?
- Yes
- No

Figure 32 Scenario 2 of how to answer questions.
Does your vertigo last between seconds to a few minutes?

○ Yes
○ No

Figure 33 First question of this section
Does your vertigo last between 20 minutes and an hour?

○ Yes
○ No

Figure 34 Second question of this section
Does your vertigo last more than an hour but less than 12 hours?

○ Yes
○ No

Figure 35 Third question of this section
Does your vertigo last more than 12 hours but not more than 1 day?

- Yes
- No
Figure 37 5th question of this section
Is your vertigo with you 24 hours a day, never stopping.

○ Yes
○ No
Figure 39 Summary of this section

**Summary**

Your vertigo lasts more than 12 hours but not more than a day; more than a day; and 24 hours a day, never stopping.

Is this correct? If not, repeat section.
The fourth section includes 12 questions asking about patients’ symptom triggering method, which shows in Figure 40. Figure 41-52 show the questions in this section. Figure 53 is the summary of patients’ answers.
Is your vertigo typically made worse or triggered by sitting up or standing up?

○ Yes
○ No

Figure 41 First question of this section
Is your vertigo typically made worse or triggered by reaching or bending?

○ Yes
○ No

Figure 42 Second question of this section
Is your vertigo typically made worse or triggered by tilting your head back?

- Yes
- No

Figure 43 Third question of this section
Is your vertigo typically made worse or triggered by lying down in bed?

Notice: The next question is about rolling over in the bed. Please distinguish it with lying down in the bed.

○ Yes
○ No
Is your vertigo typically made worse or triggered by rolling over or turning in bed?

○ Yes
○ No

Figure 45 5\textsuperscript{th} question of this section
Is your vertigo typically made worse or triggered by turning your head while walking?

- Yes
- No

Figure 46 6th question of this section
Is your vertigo typically made worse or triggered by supermarket aisles?

- Yes
- No
Is your vertigo typically made worse or triggered by seeing things go past over and over like scrolling on your phone or watching a train pass?

- Yes
- No
Is your vertigo typically made worse or triggered by loud sounds?

- Yes
- No

Figure 49 9th question of this section
Is your vertigo typically made worse or triggered by automobile rides?

- Yes
- No

Figure 50 10th question of this section
Is your vertigo typically made worse or triggered by walking on uneven ground?

- Yes
- No

Figure 51: 11th question of this section
Is your vertigo typically made worse or triggered by driving a car at night? (especially in dark surroundings)

○ Yes
○ No

Figure 52 12th question of this section
Summary

Your vertigo typically was made worse or triggered by: sitting up or standing up; turning your head while walking; supermarket aisles; automobile rides; and driving a car at night.

Is this correct? If not, repeat section.
The fifth section asks about patients’ ear conditions, which has five questions. The instruction page shows in Figure 54. Figure 55-59 show the questions in this section. Figure 60 is the summary page of this section.
Do you have ear pain?

- Yes
- No

Figure 55 First question of this section
Figure 56 Second question of this section

Do you get an ear infection at least once a year?

- Yes
- No
Do you hear a ringing, buzzing, hissing, or other noise in ONE EAR ONLY or ONE EAR MORE THAN THE OTHER?

○ Yes
○ No

Figure 57 Third question of this section
Do you hear a ringing, buzzing, hissing, or other noise in BOTH EARS EQUALLY (or in your head)?

- Yes
- No
Do you have ringing, buzzing, and hissing in your ears

○ Yes
○ No

Figure 59 5\textsuperscript{th} question of this section
Summary

You get an ear infection at least once a year. You hear a ringing, buzzing, hissing or other noise in one ear only or in one ear more than the other, and it is.

Is this correct? If not, repeat section.
The seventh section includes 10 questions. The topic of this section is headache-associated questions. Figure 61 is the introduction of this section. Figure 62-71 are the questions. The summary of answers is shown in Figure 72.

![This is a section about headaches and related symptoms.]

There are 10 questions in this section.

Figure 61 Introduction of seventh section
Have you had a total of 5 or more bad headaches in your lifetime?

- Yes
- No

Figure 62 The first question of this section
Have you ever had a headache that throbs or pulses?

- Yes
- No

Figure 63 The second question of this section
Have you ever had increased sensitivity to light with a headache?

- Yes
- No

Figure 64 The third question of this section
Have you ever had increased sensitivity to sounds with a headache?

○ Yes
○ No
Have you ever had increased sensitivity to smells with a headache?

○ Yes
○ No

Figure 66 The 5th question of this section
Have you ever had nausea or vomiting with a headache?

- Yes
- No

Figure 67 The 6th question of this section
Figure 68 The 7th question of this section

Have you ever had your vertigo associated with a headache?

- Yes
- No

PREVIOUS    NEXT
Have you ever been diagnosed with migraine headaches?

- Yes
- No

Figure 69 The 8th question of this section
Have you ever been diagnosed with ocular or retinal migraines (means a temporary change in vision that may come with a headache)?

- Yes
- No

Figure 70 The 9th question of this section
Figure 71 The 10th question of this section
You got carsick as a child. You have had more than 5 bad headaches in your lifetime. With a headache, you have had increased sensitivity to sounds and nausea or vomiting. You have been diagnosed with migraine headaches.

Is this correct? If not, repeat section.
Figure 73 shows the last section asks about patients’ medical history and the number of questions is two. Figures 74 and 75 ask questions about the patients' medical history. Figure 76 is the ending page of the survey. The database recording answers will be generated once patients press “done” button.
Figure 74 The 1st question of this section
Have you had chemotherapy?

- Yes
- No
You have completed the patient survey.

Thank you for your cooperation and please return the tablet to the study coordinator.