**Benefits Claims Decision Support System (BCDSS)**

**00031AA Predictive Model of the Ear – Hearing Loss**

**January 2017**



**Version 1.0**

Contents

[1 Executive Summary 1](#_Toc474221399)

[2 Overview 1](#_Toc474221400)

[2.1 Introduction 1](#_Toc474221401)

[2.2 Supporting Claims Data 2](#_Toc474221402)

[2.3 BCDSS Model Objectives and Basis for Ear Model Disaggregation 2](#_Toc474221403)

[3 General Hearing Loss Model Results 4](#_Toc474221404)

[3.1 General Hearing Loss Model Performance Results 4](#_Toc474221405)

[3.2 Selection of General Hearing Loss Model (Case No. 4) 6](#_Toc474221406)

[4 Tinnitus Model Results 7](#_Toc474221407)

[4.1 Tinnitus Model Performance Analysis 7](#_Toc474221408)

[5 Results for Both GHL and Tinnitus Models 8](#_Toc474221409)

[6 BCDSS Modeling - Business Process Integration 8](#_Toc474221410)

[Appendix A: Technical Specifications 10](#_Toc474221411)

[A.1 Claim Data Computations 10](#_Toc474221412)

[A.2 Claimant Data Calculations 11](#_Toc474221413)

[A.3 General Hearing Loss DDM 12](#_Toc474221414)

[Appendix B: BCDSS Model Architecture 14](#_Toc474221415)

[B.1 BCDSS Pattern-Matching Architecture 17](#_Toc474221416)

[Appendix C: Summary of MITRE Corporation’s Methodology 20](#_Toc474221417)

# Executive Summary

The Benefits Claims Decision Support System (BCDSS) team was tasked by the then acting Under-Secretary for Benefits, Tom Murphy, to decompose the aggregate statistical model developed to statistically adjudicate supplemental claims for increased compensation for Ear disabilities. The previous aggregate approach produced a Combined Disability Determination (CDD) for a suite of ear-related disabilities, as opposed to a single disability determination for a specific ear-related disability. The BCDSS team’s analysis of model performance established that the very high percentage of Tinnitus Claims for Increase either alone or in combination with General Hearing Loss (GHL) Claims skewed the results of the Ear Model. This resulted from the fact that Tinnitus original claims are rated overwhelmingly at 10 percent and almost never granted increases in subsequent claims whereas the GHL outcomes are more randomly distributed. The team proceeded to decompose the aggregate Ear Model into separate models for GHL and Tinnitus.

The decomposed models for GHL and Tinnitus increased the volume of supplemental claims rated over the original aggregate Ear Model by 13 percent. The GHL Model also demonstrated significant accuracy in predicting the rating decisions for the 72 percent of claims where the decision is Confirmed and Continued (C&C). However, the GHL model is only able to accurately predict the rating for 13 percent of claims where an increase in rating is granted.

This document elaborates on the process by which the GHL model was developed and its outcomes and provides observations and recommendations on a path forward for future models.

# Overview

## Introduction

The ***00031AB Comparative Performance Analysis of Ear Models***  describes the objectives and performance results of the Benefits Claims Decision Support System (BCDSS) General Hearing Loss Model (BCDSS Model No. 1.1) and Tinnitus Model (BCDSS Model No. 1.2).

The General Hearing Loss and Tinnitus models represent two models within a portfolio of predictive models for BCDSS. BCDSS is a proof-of-concept project jointly sponsored by the Department of Veterans Affairs (VA) Center for Innovation (VACI) and the Veterans Benefits Administration (VBA) to demonstrate the feasibility of using automation and predictive models to calculate the disability rating for certain types of claims for disability compensation benefits. The BCDSS project includes both the development of candidate models for rating claims and a software platform which provides the functionality and environment for the models to run against Veteran and Claims data in order to produce a model recommended rating.

## Supporting Claims Data

Data used to support model development, testing, and comparative analysis were derived from the same data set used by MITRE Corporation[[1]](#footnote-1) and the BCDSS team to develop the initial aggregate ear and knee related models. This data encompassed claims received by VBA between 2005 through 2014.[[2]](#footnote-2) The distribution of claims that were deemed “eligible” for use in modeling are provided in Table 1.

Table : Distribution of Claims Deemed Eligible for Modeling



Claims are deemed eligible based on whether the claim is a supplemental claim (End Product 02X), the claimant is making a relevant contention (defined as Contention Classification Codes 2200, 3140, 3150, 4700, 5710, or 6850), and VBA previously promulgated a service connected prior decision for the specific diagnosis code (Codes 6100 for General Hearing Loss, or 6260 for Tinnitus).

## BCDSS Model Objectives and Basis for Ear Model Disaggregation

The objective for all BCDSS models is to accurately determine the rating for the maximum number of specified disability conditions[[3]](#footnote-3) (e.g., hearing loss) contained within supplemental claims for increased disability compensation. There are two distinct metrics for this objective:

* **Accuracy** refers to the model calculating the same rating as that determined by the Rating Veteran Service Representative (RVSR) through VBA’s traditional claims adjudication process.
* **Through-put** refers to the supplemental claims containing at least one specified condition for which the rating can be calculated by the model (stated as a percentage).

The analysis completed by MITRE[[4]](#footnote-4) demonstrated that accuracy rates of greater than 90 percent were possible for aggregate ear-related conditions in just over 50 percent of eligible supplemental claims. Initial BCDSS hearing model (BCDSS Model 1.0) development efforts for aggregate ear-related conditions achieved accuracy rates of 94.38 percent for 60.81 percent of eligible supplemental claims. Per the direction of the Sponsor the initial Hearing Loss Model 1.0 was developed using the hearing related codes referenced in footnote 2. )

A performance analysis of the initial BCDSS Ear Model 1.0 results (Figure 1) revealed three key findings:

Figure 1: Distribution of Modeling Results by Diagnostic Code

1. 93 percent of all claims eligible to be rated using BCDSS Hearing Model No. 1.0 were rated using the General Hearing Loss (6100) or Tinnitus (6260) diagnosis codes, i.e., 23 percent of the claims contained hearing loss code ONLY, 14 percent contained Tinnitus code ONLY, and 57 percent contained claims for both General Hearing Loss and Tinnitus codes.
2. Tinnitus, once substantiated and rated, must receive a rating of 10 percent. As a result, the rating for supplemental claims where the claimant has previously received a 10 percent rating for Tinnitus will remain unchanged at 10 percent.
3. Because of the dominance and predictability of Tinnitus in the population of hearing eligible supplemental request for increase claims, VACI requested that the team proceed to decompose the model into its constituent diagnostic codes, then redevelop and re-evaluate the resulting models to determine if separate models provide improved results for accuracy and throughput. The specific objective for the General Hearing Loss model is to demonstrate that ***a condition-specific General Hearing Loss model can yield an accuracy rate equal to or greater than 95 percent across more than 60 percent of eligible claims.***

# General Hearing Loss Model Results

## General Hearing Loss Model Performance Results

The BCDSS Team developed three General Hearing Loss-specific models as part of its efforts to obtain maximum model performance relative to accuracy, throughput, and pattern rate of use (or confidence level). This “optimization” process is described in greater detail in [Appendix B](#_Appendix_B:_BCDSS).

1. ***Co-occurring Conditions Model***: A model developed using predictive characteristics (or features) for service-related conditions that co-occur most frequently with a service-related disability for general hearing loss. Analysis determined that the five most frequent co-occurring conditions were Tinnitus (Code 6260), Post-Traumatic Stress Disorder (Code 9411), Diabetes mellitus (Code 7913), Lumbosacral or Cervical Strain (Code 5237), and Degenerative Arthritis of the Spine (Code 5242). Three features were selected and calculated for each co-occurring condition based on the previously successful model development for the aggregate Ear model. These features were found to provide predictive value sufficient capture the amount of change (or magnitude) and the trajectory of the change over time. These include the:
   1. Rating determination immediately preceding the date of claim for the target claim
   2. Time period that had elapsed between that rating determination and the date of claim for the target claim
   3. Magnitude of the change of that rating relative to any previous rating for the condition

This set of 15 features were combined with the same features originally calculated for general hearing loss (as well as the age of the claimant at the time of filing) and applied to the development data set. Consistent with the feature-set optimization process described in [Appendix B](#_Appendix_B:_BCDSS), four separate “cases” were developed to establish how individual features, and the pattern set as a whole, performed for through-put and accuracy, as well as rate of use. In cases 1 and 2, specific features (the direction and magnitude of rating change, or time period between ratings) were omitted. In cases 3 and 4, the level of resolution for time periods were increased (e.g., 3.45, 3.5 and 3 years). Initial modeling results are provided in Table 2 below.

Table : Initial Modeling Results



1. ***Exceptions-Based Co-occurring Conditions Model:*** A second model was developed in an attempt to accurately identify ***ONLY THOSE CLAIMS WHERE THE RATING WAS UNCHANGED*** by the RVSR. Preliminary analysis conducted during development of the Co-occurring Conditions Model showed that a change in the rating for general hearing loss in supplemental claims was granted very infrequently. The rating was Confirmed and Continued (C&C)—or remained unchanged—in 72 percent of the 195,780 eligible claims (see Section 2.2). The performance results from the Co-Occurring Conditions Model that patterns comprised of co-occurring features calculated the rating for general hearing loss at a maximum accuracy rate of 78 percent. This accuracy rate was well below the established performance threshold of 90 percent. Based on this finding, the Team re-focused the modeling efforts on identifying the exceptions (those claims where the rating changed).

For the 72 percent of eligible claims which contained a diagnosis of General Hearing Loss, the Exceptions-based Co-occurring Conditions Model accurately forecast the C&C rating in **140,967** of the **141,105** claims. This is an accuracy rate of 99.9 percent, with a through-put of 72 percent for C&C claims. ***However, the model was designed with this objective, and consequently did not address claims where a change in rating was promulgated by the RVSR.***

1. ***Composite General Hearing Loss Model:*** Finally, the BCDSS Team developed a composite model that incorporated the features from the Ear Model (1.0) with those of the Co-occurring conditions. This feature set is presented in Table 3 below. NOTE: The definition of each feature, including the distinction between a “decision” and “rating,” is provided in [Appendix A](#_Appendix_A:_Technical).

Table : Composite Model Features



Given the previous limitations of the Exceptions-based Co-occurring Condition Model, the Composite General Hearing Loss Model was optimized to accurately identify the rating for maximum number of claims where the rating changed, while still satisfying the overall throughput and accuracy performance standards established by the original Ear Model (i.e., 50 percent throughput with an accuracy rate of 90 percent). Variations in the model - or “Cases” - and its feature set yielded range of results, but generally delivering accurate ratings for changes in General Hearing Loss rating determinations in approximately **13 percent** of the sample (while still rating C&C claims at **99 percent accuracy**). The performance metrics for the four final Cases of the model are presented in Table 4.

Table : Performance Metrics



## Selection of General Hearing Loss Model (Case No. 4)

The BCDSS Team selected Case 4 as the optimal variant of the General Hearing Loss model based on general accuracy and throughput characteristics, as well as the model’s performance in identifying and calculating ratings for non-C&C claims. **54,675** eligible claims were determined to merit a change in rating during the time period. The selected Model (Case 4) identified **18,932** of these claims (or 34.6 percent) and accurately calculated the rating for **7,313** of these claims. The selected model achieved this while also sustaining accuracy and throughput performance levels above the original 50 percent/90 percent specifications for the aggregate Ear Model. As will be described later in this document, the selected model, when combined with the revised Tinnitus Model, significantly increased overall throughput while maintaining accuracy levels above 94 percent.

# Tinnitus Model Results

The Tinnitus BCDSS Model No. 1.2 is designed to first establish eligibility. Upon eligibility confirmation, the model simply assigns the issue within the claim a rating of 10 percent. This is consistent with the rating schedule, which specifies:

6260 Tinnitus - Recurrent 10 Percent

* Note (1): A separate evaluation for tinnitus may be combined with an evaluation under diagnostic codes 6100, 6200, 6204, or other diagnostic code, except when tinnitus supports an evaluation under one of those diagnostic codes.
* Note (2): Assign only a single evaluation for recurrent tinnitus, whether the sound is perceived in one ear, both ears, or in the head.
* Note (3): Do not evaluate objective tinnitus (in which the sound is audible to other people and has a definable cause that may or may not be pathologic) under this diagnostic code, but evaluate it as part of any underlying condition causing it.

Similarly, the VBA rating calculator generates the following in its award explanation:

We have assigned a 10 percent evaluation based on recurrent tinnitus. A single evaluation for recurrent tinnitus is assigned, whether the sound is perceived in one ear, both ears, or in the head. This is the highest schedular evaluation allowed under the law for tinnitus, recurrent.

## Tinnitus Model Performance Analysis

The team tested the Tinnitus model to verify that the business logic, when applied to eligible claims, performed within established specifications. Application of the model had an accuracy rate of **99.64 percent** for **100 percent** of the **243,745** eligible claims.

The 883 claims containing a Tinnitus diagnosis that differed from the specified business rule (i.e., the assigned rating was NOT 10 percent) broke down as follows:

* 880 were assigned a rating of 0 percent.
* Two were assigned a rating of 20 percent.
* One was assigned a rating of 100 percent.

The Office of Compensation Service suggested that the 0 percent ratings may well have been assigned in error or resulted from a 0 percent evaluation, i.e., the exam revealed that the claimant did not suffer from Tinnitus.

# Results for Both GHL and Tinnitus Models

The combined number of eligible claims for the tested dataset (for both Tinnitus and General Hearing Loss) was [291,885](#Table1). This data set included 147,640 claims that contained both diagnostic codes, 48,140 claims containing general hearing loss only diagnosis codes, and 96,105 claims containing only Tinnitus diagnosis codes. The General Hearing Loss and Tinnitus models collectively addressed a total of 213,636 claims (or 73.19 percent) at a combined accuracy rate of 94.24 percent. This specific distribution and performance levels, by type of condition, is presented in Table 5.

Table : Diagnosis Codes within Claims



# BCDSS Modeling - Business Process Integration

VBA’s current claims adjudication process requires claimants to file a supplemental claim for increased disability compensation when they believe they are eligible based on a change in condition, information, or circumstances. Similar to the original claim for benefits, the claimant must present medical evidence sufficient to substantiate the condition, if they believe the condition has worsened and to what degree. After the veteran submits the necessary evidence, appropriately credentialed VBA staff review the claim and authorize promulgating the decision.

Statistical claims adjudication uses historical data captured in the filing patterns and substantiating diagnostic information for millions of claimants to establish a statistical basis for predicting the rating for a claim containing similar factual. Claimants must provide evidence of service connection and an initial claim with a fully substantiated rating to file supplemental claims. As a result, BCDSS models currently address only supplemental claims (EP-02X) where the claimant has received at least one prior service-connected rating for the modeled condition. In the future, Statistical Adjudication where model performance is acceptable could allow VBA to calculate the rating of a claim without requesting additional substantiating medical evidence.

The Office of Compensation Service originally envisioned using the BCDSS-calculated rating to provide a claimant the opportunity to accept his/her statistically adjudicated rating (where an increase might be granted) or electing to have his/her claim adjudicated through the traditional process. Analysis of the claims data by the BCDSS Team revealed that VBA grants only 40 percent of hearing-related supplemental claims for increased compensation. Of these, 27 percent are for general hearing loss and the balance are for hearing conditions other than Tinnitus.

Although rating claims using statistical adjudication would require statutory and regulatory changes, the potential benefits are enormous. Veterans would see almost instantaneous feedback on their claim and accelerated adjudication where there is a positive decision. The VA would also realize efficiencies through reduced workload. However, additional analysis is necessary to adequately quantify these efficiencies, and to ensure models can be applied effectively at a scale and operational pace commensurate with current VBA claims adjudication operations. Hence, VACI and VBA identified four potential uses for BCDSS in its current form:

1. Delivering an automated rating to the RVSR (to inform the final claim rating)–In this case, the rating would be calculated immediately following establishment and passed to the RVSR following claim development via the BCDSS user-interface.
2. Delivering an automated rating to the Decision Review Officer (DRO) or Senior Veteran Service Representative (VSR) during claim authorization (to assist in the review and verification of the manually calculated rating)–In this case, the rating would be calculated immediately following establishment and passed to the Authorizer following award generation by the responsible VSR.
3. Delivering large sets of BCDSS statistically adjudicated claim decisions to VBA quality assurance staff as means to evaluate the accuracy of claim adjudication decisions, or trends in accuracy across multiple Regional Offices.
4. Providing BCDSS access to RVSRs to support training.

# Appendix A: Technical Specifications

BCDSS Model requirements stipulate that the three core components be defined for each BCDSS model:

1. The Decision Determination Matrix (DDM)
2. The Target Claim Feature-Set Data and Computational Specifications
3. The Claimant Feature-Set Data and Computational Specifications

The feature set for the General Hearing Loss Model is provided in Table 6 below.

Table : General Hearing Loss Model Feature Set



# A.1 Claim Data Computations

SQL-1: To obtain for the target claim:

* Veteran ID (ptcpnt\_vet\_id)
* Claim ID (bnft\_claim\_id)
* End Product (end\_prdct\_type\_cd)
* Count of Ear contentions (contention\_cound)

SELECT ptcpnt\_vet\_id, bnft\_claim\_id, date\_of\_claim, end\_prdct\_type\_cd, count(cntntn\_clsfn\_ id) as contention\_count

FROM ahXXXX\_rating\_corp\_claim

WHERE end\_prdct\_type\_cd like ‘02%’ AND

cntntn\_clsfn\_id in (‘2200’, ‘3140’, ‘3150’, ‘4700’)

GROUP BY ptcpnt\_vet\_id, bnft\_claim\_id, date\_of\_claim, end\_prdct\_type\_cd

SQL-2: To obtain Count of Claims with Target Contentions with target claim:

* Claim Count for Code 2200
* Claim Count for Code 3140
* Claim Count for Code 3150
* Claim Count for Code 4700

SELECT count(unique bnft\_claim\_id) as [count for target code]

FROM ahXXXX\_rating\_corp\_claim

WHERE ptcpnt\_vet\_id=[target claim.ptcpnt\_vet\_id] and Date\_of\_Claim<[target claim.date\_of\_claim] and

Ctntnt\_clsfn\_id=[target Code]

SQL-3: To obtain current General Hearing Loss rating (Diagnosis Code 6100) for claim:

SELECT prcnt\_nbr as [Rating]

FROM ahXXXX\_rating\_decision

WHERE ptcpnt\_vet\_id=[target claim.ptcpnt\_vet\_id] AND diagnosis\_code=’6100’ AND min(prfil\_dt)>=[target claim.date\_of\_claim]

AND narrative\_codesheet\_ind=’N’ AND (ahXXXX\_rating\_decision.dsblty\_decn\_type\_cd= 'SVCCONNCTED' OR ahXXXX\_rating\_decision.dsblty\_decn\_type\_cd='1151GRANTED')

GROUP BY ptcpnt\_vet\_id, max(begin\_dt)

# A.2 Claimant Data Calculations

**Claimant Age:** Year from [target claim.date\_of\_claim]-Veteran Date of Birth (dob) rounded to nearest 10 year interval (e.g., 20, 30, 40, etc.) obtained from:

SQL-1: SELECT dob from ahXXXX\_person WHERE ptcpnt\_vet\_id=[target claim.ptcpnt\_vet\_id]

**Prior Rating (for General Hear Loss):**

SQL-2: SELECT prcnt\_nbr as [prior rating] FROM ahXXXX\_rating\_decision

WHERE ptcpnt\_vet\_id=[target claim.ptcpnt\_vet\_id] AND diagnosis\_code=’6100’ AND max(prfil\_dt)<[target claim.date\_of\_claim]

AND narrative\_codesheet\_ind=’N’ AND (ahXXXX\_rating\_decision.dsblty\_decn\_type\_cd= 'SVCCONNCTED' OR ahXXXX\_rating\_decision.dsblty\_decn\_type\_cd='1151GRANTED')

GROUP BY ptcpnt\_vet\_id, max(begin\_dt)

**Prior Rating Age: [**target claim.date\_of\_claim]-[prior rating profile date] / 365 (Rounded to numeral)

SQL-3: SELECT prfil\_dt as [prior rating profile date] FROM ahXXXX\_rating\_decision

WHERE ptcpnt\_vet\_id=[target claim.ptcpnt\_vet\_id] AND diagnosis\_code=’6100’ AND max(prfil\_dt)<[target claim.date\_of\_claim]

AND narrative\_codesheet\_ind=’N’ AND (ahXXXX\_rating\_decision.dsblty\_decn\_type\_cd= 'SVCCONNCTED' OR ahXXXX\_rating\_decision.dsblty\_decn\_type\_cd='1151GRANTED')

GROUP BY ptcpnt\_vet\_id, max(begin\_dt)

**Prior Ratings for Related Diagnosis Codes (6260, 5237, 7913, 9411, 5242):**

SQL-4: SELECT prcnt\_nbr as [Target Code rating] FROM ahXXXX\_rating\_decision

WHERE ptcpnt\_vet\_id=[target claim.ptcpnt\_vet\_id] AND diagnosis\_code=[target code] AND max(prfil\_dt)<[target claim.date\_of\_claim]

AND narrative\_codesheet\_ind=’N’ AND (ahXXXX\_rating\_decision.dsblty\_decn\_type\_cd= 'SVCCONNCTED' OR ahXXXX\_rating\_decision.dsblty\_decn\_type\_cd='1151GRANTED')

GROUP BY ptcpnt\_vet\_id, max(begin\_dt)

**Count of Prior Ear Related Decisions by Diagnosis Code (6100, 6200, 6201, 6202, 6204, 6205, 6207, 6209, 6210, 6260)**

SQL-6: SELECT

count(distinct case when ahXXXX\_rating\_decision.diagnosis\_code='6100' then null else ahXXXX\_rating\_decision.diagnosis\_code end) as A6100,

count(distinct case when ahXXXX\_rating\_decision.diagnosis\_code='6200' then null else ahXXXX\_rating\_decision.diagnosis\_code end) as A6200,

count(distinct case when ahXXXX\_rating\_decision.diagnosis\_code='6201 then null else ahXXXX\_rating\_decision.diagnosis\_code end) as A6201,

count(distinct case when ahXXXX\_rating\_decision.diagnosis\_code='6202' then null else ahXXXX\_rating\_decision.diagnosis\_code end) as A6202,

count(distinct case when ahXXXX\_rating\_decision.diagnosis\_code='6204' then null else ahXXXX\_rating\_decision.diagnosis\_code end) as A6204,

count(distinct case when ahXXXX\_rating\_decision.diagnosis\_code='6205' then null else ahXXXX\_rating\_decision.diagnosis\_code end) as A6205,

count(distinct case when ahXXXX\_rating\_decision.diagnosis\_code='6207' then null else ahXXXX\_rating\_decision.diagnosis\_code end) as A6207,

count(distinct case when ahXXXX\_rating\_decision.diagnosis\_code='6209' then null else ahXXXX\_rating\_decision.diagnosis\_code end) as A6209,

count(distinct case when ahXXXX\_rating\_decision.diagnosis\_code='6210 then null else ahXXXX\_rating\_decision.diagnosis\_code end) as A6210,

count(distinct case when ahXXXX\_rating\_decision.diagnosis\_code='6211' then null else ahXXXX\_rating\_decision.diagnosis\_code end) as A6211,

count(distinct case when ahXXXX\_rating\_decision.diagnosis\_code='6260’ then null else ahXXXX\_rating\_decision.diagnosis\_code end) as A6260

from ahXXXX\_rating\_decision where ptcpnt\_vet\_id=[target ptcpnt\_vet\_id] and [prfil\_dt]<[target date\_of\_claim] and AND narrative\_codesheet\_ind=’N’ AND (ahXXXX\_rating\_decision.dsblty\_decn\_type\_cd= 'SVCCONNCTED' OR ahXXXX\_rating\_decision.dsblty\_decn\_type\_cd='1151GRANTED')

GROUP BY ptcpnt\_vet\_id, max(begin\_dt)

# A.3 General Hearing Loss DDM

The data determination matrix for General Hearing loss is composed of 32,547 patterns, included in 2 tables (one describes the patterns, the other is an index of the patterns that should be used to obtain the optimal results). These two tables will be provided uploaded to GITHUB. The structure of each is provided below:

**GHL-DDM Index**

1. Pattern\_ID (unique Identifier)
2. Rating (Calculated Rating for Code 6100)
3. Use\_Rate (Number of Occurrences)
4. PACC (Accuracy across the instances of use)

**GHL-DDM**

|  |  |
| --- | --- |
| **FEATURE** | |
| Pattern\_ID | A6204 |
| Age | A6205 |
| Ear\_CNTNT\_Count | A6207 |
| Prior\_NBR | A6209 |
| Prior\_NBR\_Age | A6210 |
| C2200 | A6211 |
| C3140 | A6260 |
| C3150 | D6260\_PRCNT\_NBR |
| C4700 | D9411\_PRCNT\_NBR |
| A6100 | D7913\_PRCNT\_NBR |
| A6200 | D5237\_PRCNT\_NBR |
| A6201 | D5242\_PRCNT\_NBR |
| A6202 |  |

# Appendix B: BCDSS Model Architecture

The BCDSS Team evaluated MITRE’s analytical methodology for statistically adjudicating claims for scalability and the ease with which the classifiers could be automated. The BCDSS Team’s evaluation determined that a simpler approach using a pattern-matching architecture could streamline rating calculation and prediction, while providing enhanced scalability. This approach retains the ability to apply Ordinal Logistic Regression (OLR) or other advanced statistical algorithms should such analytical capabilities become necessary to satisfy performance parameters. The BCDSS Team based this hypothesis on the following:

* A cursory analysis of supplemental claims for increased compensation containing at least one ear-related contention indicated that most claimants file an average of less than two such claims. The team concluded event-specific analysis might yield just as accurate a result given the limited number of relevant events contained within most claimants’ adjudication history.
* Establishing an architecture that can use and modify a claimant’s adjudicatory history, as well as training data to derive a single Combined Disability Determination (CDD), seemed unnecessarily cumbersome. It might also present significant performance challenges when there are a high numbers of users and transaction requests. The final design concept needs to be repeatable, efficient, and highly scalable.
* MITRE applied machine-learning algorithms to establish predictive adjudicatory patterns from historical training and test data and then applied the pattern against target claims to calculate the likely CDD. Current computer processing speeds can execute pattern-matching routines far faster and with greater efficiency than machine-learning algorithms, even with inventories of more than 100,000 patterns.
* A basic modeling architecture that relies on matching a fixed set of claim attributes for a specified claim that includes the features contained within an inventory of patterns is extremely scalable. It can also exclude invalid patterns. This would be far more challenging to accomplish with statistical adjudication models.

Based on this, the team developed and tested a more streamlined and scalable design concept. This design concept assumes that the attributes related to specific conditions and the decisions of VA personnel who adjudicate claims for those conditions form repeated and predictable patterns. These patterns are captured within the values of a predefined suite of predictive characteristics or features. A sample pattern inventory is shown in Figure 5:



Figure : Sample Pattern Inventory

The BCDSS Team tested its hypothesis with the same data that MITRE used to conduct its analysis and the same feature set:

1. The team extracted claims that included one or more ear-related contentions from the initial set of 1.2 million claims.
2. The data were grouped based on unique combinations of values for the set of features.
3. The patterns were numbered and aligned with the resulting CDD for the ear (the calculated current CDDs were rounded to the nearest 10 percent or quantified based on MITRE’s analysis).
4. The team executed an automated routine to match the same set of claims with the pattern inventory to determine whether the logic was self-repeating (i.e., the claims used to identify the patterns could again be matched to provide an accurate result).
5. The initial test indicated that fewer than 65,000 patterns were required to determine the CDD at accuracy rates above 90 percent. Based on these initial results, the team adopted the pattern-matching architecture for preliminary development and testing purposes.

The BCDSS Team used a four phase technical approach (Figure 6 below) consistent with previous models.



Figure : BCDSS Model Development Approach

* **During Phase 1,** the team extracted and segmented data to ensure sufficient claimants with relevant claimed contentions and adjudication decisions are included to form distinct development and test data sets. The team then synthesized the data within the development data set to form a high-resolution feature set (i.e., values were calculated to the highest level of precision in the underlying data: in this case, time periods were calculated in days where defined by starting and end dates).
* **During Phase 2,** the team aggregated the development data to form unique, numbered patterns, and aligned the patterns with the corresponding CDDs. Sensitivity and optimization analyses were conducted on each feature to identify the relevant impact of the feature on accuracy, and the extent to which the feature could be aggregated to increase the rate at which a pattern is re-used across test data without compromising the model’s predictive accuracy. For example, storing the period between the claim date and last CDD in years rather than days.
* **During Phase 3,** the team optimized the model to maximize “rate of use” within the context of target through-put and accuracy performance parameters. The confidence level of pattern sets are a function of repetition (or “rate of use”) relative to the size of the sample (number of claims eligible for modeling).
* **During Phase 4,** the team tested the model against an entirely separate set of verification data to verify that the performance results initially obtained during optimization were replicated, and confidence levels were sustained. The government was also provided the model to allow its analysts to conduct independent analyses of the models.

## B.1 BCDSS Pattern-Matching Architecture

The BCDSS pattern-matching modeling architecture relies on replicated fact patterns for similar conditions and claim adjudication histories, and the consistent promulgation of ratings for such fact patterns by VA employees to reliably predict ratings (without subjecting the claim to additional substantiation and human evaluation). A relatively small number of claim attributes and established facts contained in the claimant’s filing history comprise these patterns. The rating of a specific condition within the target claim is obtained by establishing the fact pattern of the claim using a predefined set of attributes, and matching it to a specific historical pattern and the rating that most often results. This BCDSS pattern-matching architecture contains three major components:

1. A set of claim-specific features and the associated specifications for calculating the values for these features. These features relate to the claim that will be subject to predictive modeling—the target claim.
2. A set of claimant-specific features and the associated specifications for calculating the values for these features. These features relate to attributes of the claimant’s filing history, including decisions contained in prior claims for similar conditions, the period of elapsed time between the target claim and prior decisions, the place of the target claim in the sequence of relevant claims in the claimant’s adjudication history, etc.
3. The Decision Determination Matrix (DDM) or set of predictive patterns and related CDDs for the modeled condition. The columns of the DDM constitute the various features as well as the CDD and their performance attributes. The rows are the unique combinations of values that constitute the predictive patterns.

The basic logic described above is depicted in Figure 7.



Figure : Pattern-Matching Logic

The BCDSS provides an automated environment in which authorized users select one or more target claims to adjudicate, one or more models to apply, and one or more output formats preferred by the user. The system then executes automated routines in accordance with the user selections to derive the results. Three inter-related modeling engines support these functions as depicted in Figure 8.



Figure : BCDSS System Workflow and Modeling Architecture

Upon user selection, the BCDSS retrieves appropriate model(s) from the model repository. The ingest engine verifies that the claim satisfies the eligibility requirement and executes the required calculations against the target claim and related claimant data to create a composite feature set. The modeling engine then conducts the pattern matching and either assigns the associated CDD (where a match is found), or identifies that the CDD could not be established. Finally, the output engine integrates descriptive information about the claim (and the CDD, where appropriate) and formats the data in accordance with the user’s selection.

Appendix C: Summary of MITRE Corporation’s Methodology

The methodology used by MITRE to complete its analyses is documented in a series of engineering notebooks (completed under Contract No. VA118A13J0421/VBA OSP COMPENSATION SERVICES CLIN 0005 IFCAP 101-J47030). MITRE’s methodology was as follows:

1. VA supplied MITRE (and later, the BCDSS Team) with a set of claims and veteran data for the purpose of the analysis. These VBA claims and rating decision data were first conditioned, which involved aligning the claims data with the corresponding decision data to produce a chronological claimant adjudication history and feature vector. This feature vector constitutes a recurring sequence of claim attributes and the trajectory of related decisions that result in common outcomes.
2. Machine-learning algorithms or “classifiers” (e.g., Random Forest, Logistic Regression, Ordinal Logistic Regression (OLR), and Auto-encoder) quantified the relative importance of a suite of common claim attributes (or features) for determining the CDD for subject conditions within supplemental claims, across samples of the above described claimant adjudication histories. Those features with greatest predictive value were subjected to testing and predictive modeling.
3. The classifiers were also applied to evaluate their relative performance in predicting the CDD of randomly selected (eligible) claims. The above-mentioned dataset, once conditioned, was segmented into a Training Set and a Test Set. Training data supplied the analytical basis for determining correlations between features to calculate the predicted CDD. Separate sample claims were then used to test the predictive capabilities.
4. Results were measured based on throughput (i.e., the ratio of the number of claims processed by the classifier to the total number of claims) and accuracy (i.e., the percentage of predicted CDDs that are equal to the CDD assigned by the responsible Rating Veterans Service Representative (RVSR).

MITRE’s work established the relative importance of specific features and identified methods that can quantify their importance. These analyses also provide performance evaluations of various classifiers MITRE applied to statistically adjudicate claims that match specific fact patterns. These patterns include the features identified as having significant predictive value, as well as prior claimant adjudication decisions, time periods between the subject claim and prior decisions, the diagnostic information used to describe those decisions, attributes of the claimant (age at the time of filing, etc.), and data within the subject claim (e.g., the contention classification codes used to describe the Veteran’s new claimed disability).

MITRE categorized its set of features as either numeric or text. Numeric values were calculated where necessary (e.g., time periods, counts, etc.), rounded to the nearest decile (e.g., 35 would be stored as 40), and otherwise stored as true/false values (e.g., whether a specific diagnostic code or contention code was used). The same true/false logic was applied to the calculation of text values (i.e., the presence of specified text strings in specified fields within the data).

1. See MITRE Study completed under Contract No. VA118A13J0421/VBA OSP COMPENSATION SERVICES CLIN 0005 IFCAP 101-J47030 and related “Statistical Adjudication Engineering Notebooks.” [↑](#footnote-ref-1)
2. Provided by VBA’s Office of Performance Analysis & Integrity (PA&I). Table names ah4929\_rating\_decision, ah4929\_rating\_corp\_claim, and ah4929\_person. [↑](#footnote-ref-2)
3. Conditions coded as Diagnosis 6100 (Hearing Loss), 6200 (Otitus Media (Chronic Ear Infection)), 6201 (Otitis Media), 6202 (Otosclerosis), 6204 (Labyrinthitis), 6205 (Meniere's Syndrome), 6207 (Loss or partial loss of Ear), 6209 (Benign growth of Ear), 6210 (Hearing Loss), 6211 (Perforated ear drum), and 6260 (Tinnitus). [↑](#footnote-ref-3)
4. See Note 1. [↑](#footnote-ref-4)