

# National Health Plan Achieved 98.2% Coding Accuracy When Using Cavo Health with Precise Word Matching AI

## EXECUTIVE SUMMARY

A major national health plan's coders achieved 98.2% risk adjustment coding accuracy when using the Precise Word Matching AI engine in the Cavo Health NLP computer-assisted coding product, according to a scientific test. The health plan's Quality Assurance process followed by an independent third-party review confirmed these results.

## BACKGROUND

Health plans are using products with Natural Language Processing (NLP) powered by machine learning for computer-assisted risk adjustment coding of medical records. The health plan previously tested two NLP coding products against its proprietary computer-assisted coding method. After these tests, the health plan concluded that its proprietary method was both more accurate for risk adjustment coding than when using these two products powered with machine-learning NLP.

After learning about a new NLP technology powered by Cavo Health's Precise Word Matching AI, the health plan decided to perform a third scientific test to determine how this new technology compared to its established risk adjustment coding method.

## STATISTICAL TEST DESIGN

### A. Research Design

Cavo Health is the third vendor that the health plan tested for computer-assisted coding. The goal of this study is to measure both coding speed and accuracy when coders use Cavo Health versus alternative traditional method processes.

### B. Sample

**Test/Practice File:** There were 60 progress notes that the coding team will use to learn the Cavo coding tool. Cavo Health was onsite to demonstrate how to use the tool effectively.

- Day One: step-by-step training with onsite support
- Day Two: hands-on coding using the 60 charts with onsite support
- Days Three & Four: hands-on coding using the 60 charts with telephonic support

**Pilot Sample:** There were 412 charts that the coding team were assessing. Of the 412 charts, 385 were outpatient and 26 were inpatient. For this audit only, the coding team used 2017 Coding Guidelines, regardless of the date of service on the chart. For the pilot, coders did not be performing complete code capture. Coders validated patient name, DOS, and provider.

### **C. Coding Process**

To ensure that the coders were not biased in their review, no coder coded a chart twice. Rather, a cross-over process was followed.

- The eight coders were broken up into four two-member teams: Team 1, Team 2, Team 3 and Team 4.
- The 412 test charts were divided into two batches of equal number: Batch A and Batch B.
- Team 1 coded with Cavo Health for Batch A, then coded Batch B in the traditional method
- Team 2 coded Batch A with the traditional method then coded Batch B with Cavo Health.
- Team 3 coded with Cavo Health for Batch B, then coded Batch A in the traditional method
- Team 4 coded Batch B of charts with the traditional method then coded Batch A with Cavo Health

### **D. Analysis Process**

After the completion of the Coding Process, the following steps were followed to generate the test results of this study:

Step One: The health plan and Cavo Health reviewed the results of the Coding Process for possible errors and reasonableness. No errors were identified. However, one coder's results were determined to be an outlier because they were inconsistent with the results of the other three coders to a degree that could not be explained. For the sake of reasonableness, the health plan and Cavo Health agreed to exclude this coder's results from the analysis of the test results.

Step Two: The health plan's Quality Assurance team coded the same charts a third time using the health plan's current methodology.

Step Three: Cavo Health identified the codes that were confirmed by the National health plan coders when using the Cavo Health system but were not confirmed by the health plan's Quality Assurance team. These codes were classified as Overturns.

Step Three: An independent third-party risk adjustment coding educator audited the Overturns. The Overturns confirmed by the third-party auditor were deducted from the total codes confirmed when using the Cavo Health system.

Step Four: Cavo Health aggregated the results of the four reviews of the test charts and submitted a report to the health plan for its review. A summary of these results is presented below.

## TEST RESULTS

### A. Overall Summary

The health plan's coders using the Precise Word Matching NLP engine in Cavo Health achieved 98.2% accuracy in their coding, as determined by a the health plan QA audit followed by a third-party review of the overturns. In addition, the coders coded 9% faster than when using the health plan's current method.

The following is an overall summary of the test results:

**Cavo Health Coding Accuracy according to an independent  
98.2% Third-Party Review**

228	National health plan QA Team Unique Overturns
9.2%	Unique Overturns as Percent of all Unique Confirmed Cavo Codes

44	National health plan QA Team Overturns Confirmed
19%	Percent of Overturns Confirmed by Third Party Review
1.8%	Unique Overturns Confirmed by independent Third-Party as Percent of all Unique Confirmed Cavo Codes

## B. Mitigating Factors

When aggregating coding results, Cavo Health decided all mitigating factors such that they would not bias the test in favor of the Cavo Health system. Cavo Health made the following decisions when calculating the results of the test:

- When the third-party reviewer audited the Overturns, he reduced the number of unique confirmed codes found with the Cavo Health product by 40. However, the third-party reviewer did not likewise audit Overturns of confirmed codes when using the health plan's method. Instead, for the purposes of the study, all coding when using the National health plan method was assumed to be correct.
  - This process potentially biases the results of the test against the Cavo Health system because the third-party auditor might have confirmed the Overturns of some of the codes confirmed when using the health plan's method.
- When determining the number of unique codes found with the Cavo Health product, the confirmed codes were de-duplicated. However, the codes found using the health plan's method were not de-duplicated. Instead, for the purposes of the study, all codes reported when using the health plan's method were assumed to be unique.
  - This process potentially biases the results of the test against the Cavo Health system. In fact, if the health plan's coders using the health plan's current method coded duplicate codes at the same rate as they recorded duplicate codes using Cavo Health, then the number of unique codes confirmed with the health plan's current method is 7% less than reported. In such a case, use of the Cavo Health system instead of the health plan's current method would have resulted in 168 more unique codes confirmed. This means the Cavo Health system would have found 1.48 additional unique codes per chart reviewed versus use of the health plan's current method.
- When auditing the Cavo Health codes that the health plan's Quality Assurance team overturned, the third-party reviewer noted when the code was present but less specific than another code that was confirmed. In this situation, the analysis assumed this code should not have been confirmed since it was not necessary to confirm an HCC.
  - This process potentially biases the results of the test against the Cavo Health system. In fact, if these codes were included, then ten more unique codes would have been confirmed with the Cavo Health system.
- The health plan's coders were familiar with the health plan's current coding methodology. In contrast, National health plan coders received three days of practice on the Cavo Health product before the start of the scientific test.
  - This Process potentially biased the results of the test against the Cavo Health system because Cavo Health reports that coders often increase their Cavo Health coding speed as they become more comfortable with the product over time.