STATE OF CONNECTICUT

AUDITORS' REPORT ON
STATE EMPLOYEE SPOUSAL
HEALTH INSURANCE ELIGIBILITY

July 9, 2015

AUDITORS OF PUBLIC ACCOUNTS
JOHN C. GERAGOSIAN   ❖   ROBERT M. WARD

Table of Contents

Executive Summary

In accordance with the provisions of Section 2-90 of the Connecticut General Statutes, we have conducted a review of potential instances in which state employees might be improperly covering ex-spouses as dependents on their state insurance plans subsequent to divorce or legal separation.

This review utilizes divorce records obtained from the Judicial Department of the State of Connecticut as well as state employee and dependent information maintained in the benefits module of Core-CT, the Connecticut state government's integrated human resources, payroll, and financial system.

To detect possible matches, the Jaro Winkler algorithm was used to compare pairs of names between judicial and Core-CT records. Jaro Winkler is a measure of similarity between two strings of text and was designed and developed for comparison of person names by the United States Census Bureau.

In a previous review conducted by the Office of the State Comptroller, a matching algorithm known as Soundex was used to complete this same comparison. Our review concluded that, when using the same data, Jaro Winkler, a more advanced matching algorithm, provided better results in detecting matches of divorced spouses who retained state insurance coverage contrary to state policy.

In addition, we examined the potential for automatically scheduling this comparison to facilitate the manual investigative work that must occur after the automated process determines which pairs of names require additional review. At the time of our review, there was no fixed interval at which this review was being completed.

Background

State employees are only permitted to cover ex-spouses as dependents on their state insurance plans under certain conditions. When they do so, they must contribute toward the cost of that coverage.

In general, entry of a divorce decree requires that the former spouse be removed as a dependent under all insurance plans (medical, dental, prescription, etc.) within 31 days from the effective date of the divorce or legal separation.

The Office of the State Comptroller (OSC) released details related to this matter in a memorandum dated May 4, 2012, entitled Need to Update Dependent Status Information/Upcoming Dependent Eligibility Audit.

The memo stated the following:

> **"1. Members divorced on or before June 30, 2012**
>
> Members who were divorced on or before June 30, 2012, and who are required to provide health benefit coverage for an ex-spouse under the laws of their home state or under the terms of a divorce decree will be allowed to continue coverage for a former spouse, subject to the following:
>
> - The member must immediately report the divorce by submitting a Form CO-1319 to their employing agency for active employees or to the Retirement Health Unit of HPBSD for retirees;
> - Coverage of the former spouse may continue for up to three years or until either party remarries, whichever comes earlier; and
> - The affected plan member must pay 100% of the cost of the individual coverage for the former spouse under the selected vendor and plan (includes both the employee and the state shares).
>
> **2. Members divorced on or after July 1, 2012**
>
> Members who divorce on or after July 1, 2012, and who are required by state law or divorce decree to provide health benefit coverage for an ex-spouse, will be allowed to continue coverage for that former spouse under the state plan, provided:
>
> - The member immediately reports the divorce by submitting a CO-1319 Form to their employing agency for active employees or to the Retirement Health Unit of HPBSD for retirees within 31 days of the date of the judgment;
> - Coverage of the former spouse may continue for up to three years or until either party remarries, whichever comes earlier; and

- The affected plan member must pay 100% of the cost of the individual coverage for the former spouse under the selected vendor and plan (includes both the employee and the state shares)."

In any circumstance other than those stated above, the employee must remove an ex-spouse from any applicable insurance plan to which the ex-spouse is enrolled. At the very least, the employee must report the divorce and have the former spouse reclassified as an ex-spouse.

To that end, OSC conducted an audit of all persons enrolled as dependents under the state employee plan as of the plan year starting July 1, 2012. In that audit, they found 223 cases in which a person was listed as a dependent on a state employee's or retiree's health coverage matching a divorced or legally separated couple on the state Judicial Department database. Of those 223 couples identified, 138 involved active employees and 85 involved retirees.

Not all of those matches necessarily represent the same people. For instance, a state employee and spouse could have the same names as another divorced couple. Unique identifying information such as a Social Security number or date of birth is not recorded in the judicial database, which is why the audit began with a comparison of names. As such, subsequent follow-up had to occur to make any concrete conclusions as to each match identified.

Upon further review of the matches identified, using other sources of information such as resident addresses and Internet searches, OSC was able to determine with certainty that 136 of the 223 matches represented the same couples. The remaining 87 matches did not have enough other information in available public records to determine with certainty whether they represent the same couples and don't simply share the same pair of names. Of the 136 confirmed matches, 98 involved active employees, and 38 involved retirees.

The matches of active employees were distributed to the Office of Labor Relations (OLR) for follow-up review. OLR distributed lists of matched employees to each agency's human resource staff to review and conduct employee interviews to determine whether those matches did indeed represent ineligible recipients. For those matches involving retirees, OSC performed the follow-up on its own, as OSC administers the state's retirement plans. In the case of the retirees, OSC immediately removed from coverage those 38 retirees who were confirmed with certainty. In the case of active employees, to date, 35 stipulated agreements have been reached. The agreements dictate the amounts to be paid to the state for time in which a former spouse was ineligibly covered. The agreements also provided either that the dependent be removed from coverage or be permitted to remain covered in accordance with the above policy, including that the employee pays the fair market value of the coverage.

Objectives, Scope, and Methodology

The office of the Auditors of Public Accounts, in accordance with Section 2-90 of the Connecticut General Statutes, is responsible for examining the performance of state entities or operations to determine their effectiveness. We conducted a review to assess how many state employees might be improperly covering ex-spouses as dependent spouses on their state insurance. This review does not constitute an audit.

We obtained a list of divorce records maintained in the state Judicial Department database as of September 8, 2014. That listing contained 157,398 unique divorce or legal separation cases with a disposition date prior to August 1, 2014.

We also obtained a list of employees covering a dependent spouse as of September 8, 2014, by querying certain database tables that are used by the benefits module of the state Core-CT system. There were 49,350 such employees.

Looking for matches between these two datasets, we used an Oracle database to perform our review. Oracle databases have a built-in package called UTL_MATCH which, according to Oracle's own documentation "…is typically used to match names, such as two first names or two last names." This package provides a function called Jaro Winkler, which uses the Jaro Winkler algorithm to calculate the similarity between two strings of text on a scale of 0 (no similarity) to 1 (exact match).

Because this is the first time that this review has been completed since 2012, we provided the detailed results of our review to the Office of the State Comptroller so that it could distribute the results to the appropriate parties for further review or perform follow-up reviews for each match identified.

Interpreting the Results of the Analysis

Comparisons Drawn

Four names were compared during this review as shown below:

#	Core-CT Database	Judicial Database
1	Employee First Name	Plaintiff First Name
2	Employee Last Name	Plaintiff Last Name
3	Dependent Spouse First Name	Defendant First Name
4	Dependent Spouse Last Name	Defendant Last Name

It should be noted that an employee can be a match with either a plaintiff or defendant in a divorce or legal separation case record on the judicial database. In one match, the employee might match up with the plaintiff and the spouse with the defendant, while in another match the employee might match up with the defendant and the spouse with the plaintiff. Matching pairs of names in either direction were considered a match in our review.

Our review identified both exact and fuzzy matches as defined below. Those matches were distributed to the Office of the State Comptroller staff to review, using resident addresses on file and Internet searches, to confirm or rule out matches as actually representing the same couples enrolled in state insurance coverage.

Exact Matches

An exact match is a perfect match between a pair of names on the judicial database and the state's Core-CT database. All four names, as listed above, on the Core-CT database match all four names of a case record on the judicial database with no spelling differences.

Fuzzy Matches

A fuzzy match is a match between a pair of names on the judicial database and the Core-CT database in which all four names, as listed above, score a Jaro Winkler similarity of 0.90 or higher, but at least one is below a perfect match of 1. Some might be identical; however, one or more of the four names compared were not identical. For instance, the following would be considered a fuzzy match in our review, even though three of the names compared were a perfect match:

#	Core-CT Database	Judicial Database	Jaro Winkler Score
1	Jack	Jack R	0.93
2	Smith	Smith	1
3	Jill	Jill	1
4	Smith	Smith	1

* Please note that this is an example with imaginary data and does not reflect an actual couple in either database.

As shown in the above example, the accidental or erroneous inclusion of a middle initial in the first name field on either database would lead to a fuzzy but not perfect match, even if the three other name comparisons were the same.

In other cases, names with multiple similar spellings might yield a fuzzy but not perfect match, such as "Louis" compared with "Luis" which has a Jaro Winkler score of 0.94.

Results of the Analysis
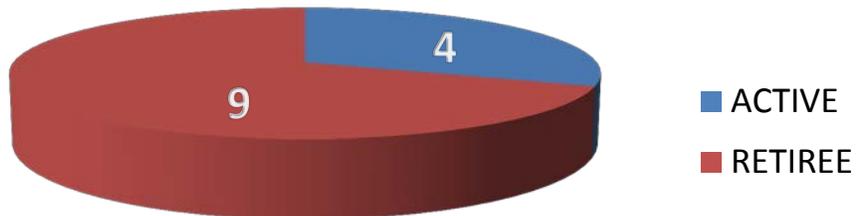
Confirmed Matches

We found 157 confirmed matches comprised of 79 active employees and 78 retirees:
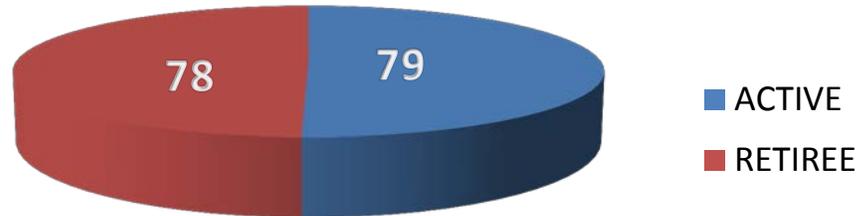
Confirmed Matches, Jaro Winkler Exclusive

Of the 157 confirmed matches, 13 matches identified by the Jaro Winkler string similarity algorithm would not be identified by the Soundex phonetic algorithm.

These are matches that would not have been identified using the Office of the State Comptroller's current matching methodology.

Confirmed Matches by Department, Active Employees
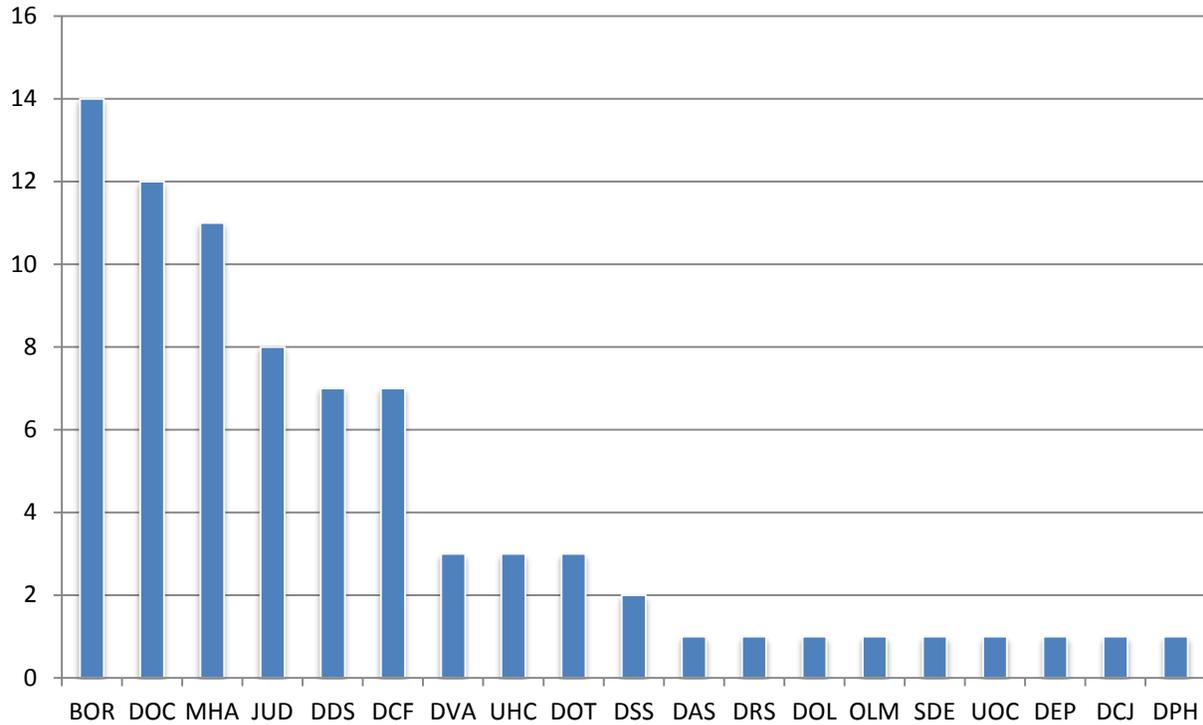
BOR	Board of Regents
DOC	Department of Correction
MHA	Mental Health and Addiction Services
JUD	Judicial Department
DDS	Department of Developmental Services
DCF	Department of Children and Families
DVA	Department of Veterans' Affairs
UHC	University of Connecticut Health Center
DOT	Department of Transportation
DSS	Department of Social Services
DAS	Department of Administrative Services
DRS	Department of Revenue Services
DOL	Department of Labor
OLM	Office of Legislative Management
SDE	State Department of Education
UOC	University of Connecticut
DEP	Department of Environmental Protection
DCJ	Department of Criminal Justice
DPH	Department of Public Health

** The above chart is limited to the confirmed active employee matches because all retirees are coded to and administered by the Office of the State Comptroller.

Auditor Analysis

The OSC review in 2012 was limited to the detection of Soundex matches between pairs of names in Core-CT and the judicial database. Soundex is a phonetic algorithm that is used for indexing names by the sound resulting from their pronunciation in English. It was developed in the first half of the 20th century, patented in 1918 and 1922. It converts every string of text to one letter followed by three numerical digits, the first letter being the first letter of the string of text and the subsequent digits encoding all remaining consonants within the string of text. Regardless of the size of the string, the remaining consonants are always encoded in only three digits. Likewise, Soundex is particularly vulnerable to missing matches resulting from typos where the person simply mistypes a name by transposing the order of the letters, as it does not gauge the similarity between strings, but evaluates how a short string of text sounds to an English speaker.

Modern improvements to the Soundex function include the Metaphone phonetic algorithm, published in 1990, and subsequent revisions to that algorithm, including Double Metaphone and Metaphone 3. The original Metaphone algorithm encodes remaining consonants in 16 symbols.

However, neither the Soundex nor Metaphone algorithms involve scoring, so they too can miss matches that are too fuzzy due to spelling errors that yield a slightly different sound from the correct spelling. The main advantage of these algorithms, in comparison to Jaro Winkler – the algorithm that we used – is that they can be applied against just one string of text, instead of both strings of text being compared. Then, the results can be indexed and matched against each other. This results in extremely fast search results and is suitable for large amounts of data; however, some reasonable matches can go unidentified. For instance, a last name of Collin would not be matched with Collins, which have Soundex values of C452 and C450, respectively. In comparison, the Jaro Winkler algorithm – which scores the similarity between two strings – identifies these two strings as 97.14 percent similar.

In our analysis of using Jaro Winkler, we found 13 confirmed matches that would not have been identified if Soundex were used. These matches would not have been found using OSC's current matching methodology.

In addition, neither OSC nor any other agency has subsequently performed this review since the initial analysis in 2012.

Because a comparison of divorce records with employee dependent records involves analyzing pairs of names held in two different databases in which the data itself is entered by different people at different times, the chances of spelling errors or other data entry errors is likely – such as including a person's middle initial in the first name field, or other variation of the data entered in each system.

As a result, relying exclusively on similar sounding pairs of names between two systems is likely to result in many missed fuzzy matches that may in fact represent the same individuals. In this case, 13 fuzzy matches – which were confirmed using address records and other means – would have gone unidentified and improper coverage would have continued indefinitely unless the OSC matching procedures were eventually strengthened.

In addition, if extract, transform, and load (ETL) processes were developed, using data from the Judicial Department as the source and Core-CT as the target system, similar names could be continuously searched as part of automated batch processes and stored in tables for review by the agencies involved and periodically monitored by the Office of the State Comptroller.

It should be noted that, while there is an obvious cost to the state for providing insurance coverage to these ineligible recipients, we were unable to estimate the dollar amount associated with that coverage. Because the state is now self-insured, the cost attributable to each state employee's health coverage varies based on the number and cost of actual claims incurred by each employee and the employee's dependents.

However, as calculated by the Office of the State Comptroller, the average annual fair market value of the medical plans offered by the state in fiscal year 2015 was $8,012 and the average annual fair market value of the dental plans was $471. Combined, the annual average is $8,483 per person. This is an approximate amount of what one would have had to pay for equivalent medical and dental coverage if they had been removed from state coverage upon divorce and had to purchase it on their own. Within our 157 matches, based on the amount of time these ex-spouses were covered subsequent to divorce, as of our review in September 2014, employees benefited by a rough estimate of about $3.6 million based on today's rates. While this is not an exact figure, because it does not consider the actual rates of the particular plans in which employees were enrolled (rather it considers the average of all plans offered) or the marginally lower costs of insurance in prior years in which coverage was held, it is a rough estimate that can provide some context as to the extent to which individuals benefited from remaining on coverage offered by the state.

Recommendations

Use of Outdated Matching Algorithm

Criteria: When status changes of dependents on state employee insurance plans occur, due to events such as divorce, employees must notify their department's human resources staff within 31 days of the event. Ex-spouses must be removed from coverage upon divorce.

This is communicated annually by open enrollment brochures distributed by the Office of the State Comptroller.

Condition: The Office of the State Comptroller currently uses an outdated matching algorithm known as Soundex to identify potentially ineligible dependent ex-spouses being covered as dependent spouses on state employee insurance plans. The office uses that algorithm to match records between the judicial and Core-CT databases.

Effect: In our review, which identified 157 confirmed matches, 13 of those matches (8%) would not have been found if we had relied on the Soundex algorithm. Those divorced employees would likely have been able to continue covering their former spouses indefinitely.

Cause: We were informed that because Soundex is the only built-in function designed for fuzzy matching available within Microsoft Access, it was the only methodology department staff involved in the review knew how to use at the time.

Recommendation: The Office of the State Comptroller should use a more advanced matching algorithm than Soundex for this analysis so that legitimate matches do not go unidentified due to the limitations of Soundex.

Agency Response: "We agree. Although we have since learned that there are more sophisticated data analysis tools than Soundex, the method used at the time of the initial audit in 2012, was best known to us and is a common feature of the data analysis programs we regularly use, such as, MS Access and MS Excel."

Frequency of Audit Procedures

Criteria: Audit procedures aimed at detecting ineligible dependents assigned to state employee insurance coverage that are not performed on an ongoing basis are unlikely to deter or detect instances of ineligible dependents in a timely manner.

The greatest chance of reducing noncompliance would be the establishment of continuous procedures that rely on automated

processes in as many ways as possible.

Condition: The Office of the State Comptroller last completed a dependent spouse eligibility audit in 2012 and at the time of our testing (September 2014), had not established any fixed interval at which such an audit would be performed.

Effect: Ineligible dependents might be identified late or not identified at all. The rate of noncompliance might be higher than it would be if audit procedures were conducted more frequently or at a defined interval.

Cause: We were informed during our testing that the department was in the process of planning to conduct another audit, but had not done so since its initial 2012 audit. In addition, the automation of some processes would rely on Core-CT staff, who are currently preparing for a PeopleSoft upgrade.

Recommendation: The Office of the State Comptroller should complete this analysis on a monthly and annual basis.

The monthly review would consist of matching divorces with a disposition date in the preceding month. The annual review would be completed upon each open enrollment and would compare the dependent spouse population with all divorces regardless of disposition date, not just those in the preceding month.

As part of this process, the Office of the State Comptroller, or a chosen service provider, should automate the following processes:

- Develop, extract, transform, and load (ETL) processes between the judicial and Core-CT databases

- Automate Jaro Winkler matching of names at the end of each month with the divorce records from the preceding month, at each monthly review, and with all divorce records at each annual review

- Load the Jaro Winkler matching results with resident address detail into the Core-CT database used by the PeopleSoft Human Capital Management system to facilitate a manual review of the matches identified

Because certain procedures cannot be automated, the Office of the State Comptroller should manually perform the following procedures:

- Review the matches identified by the matching algorithm, conduct follow-up reviews using Internet searches, and notify the Office of Labor Relations of confirmed active employee matches, and the OSC Retirement Division of confirmed retiree matches

- Mark false positives so that they are eliminated from consideration in future reviews

- Mark instances in which a match is due to a couple remarrying, or any other legitimate reason, when notified of such by the Office of Labor Relations

Agency Response: "Regarding the manual processes described:

This is part of the current process. We are establishing a reporting mechanism between our office and OLR to track the outcome of each confirmed case where judicial records indicate that an active employee is providing health benefit coverage to a former spouse following a judgment of dissolution or legal separation.

Regarding the automated processes described:

We agree. We have discussed the recommendation with the Core-CT team to determine feasibility of monthly reviews in light of other initiatives, which include the upgrade of Core-CT to version 9.2, completion of the Pension Module, and transitioning the University of Connecticut to the Core-CT Payroll system. The Core-CT team would like to meet with the Auditors of Public Accounts to discuss the details of implementing the recommendation."

Concluding Comments by the Office of the State Comptroller

"We would like to express our appreciation to the Auditors of Public Accounts for providing the initial data set of 563 names, which was used to identify some 157 ineligible spouses who were not removed from coverage following divorce or legal separation.

The Auditors of Public Accounts' data comparison produced a data set of 563 potential matches. The Judicial Department's publicly available records lack detailed identifiers, such as, social security number or birth date. For that reason it is necessary to compare initial matches against employee addresses and other details revealed by internet searches to confirm that the person identified in the judicial record is the same as the person enrolled in our health benefit program. The staff of the Healthcare Policy and Benefit Services Division reviewed each of the 563 potential matches and determined that 157 of the returned names were state employees or retirees with an enrolled spouse on the plan. Irrespective of the similarity measure algorithm used to determine potential matches, the final step to confirm and verify exact matches requires manual intervention to ensure the integrity of the data.

The following is the disposition of the records as of January 30, 2015.

Active employees:

304	Initial APA matches
95	Number of initial matches September 2014
- 3	Duplicate matches
91	Confirmed active matches sent to OLR for review October 2014
- 12	Remarried prior to initial match
79	Active confirmed matches with ex-spouses on coverage reported to APA
- 17	Remarried since initial report
62	**Final number of matches with ex-spouses on coverage. All of which have been removed from coverage with a termination date no later than 3/1/15.**

Retirees:

259	Initial APA matches
79	Number of Initially Reported Matches September 2014
- 1	Duplicate match
78	**Confirmed retiree matches, all of which OSC has removed from coverage with a termination date no later than 3/1/15.**"

Conclusion

We wish to express our appreciation for the cooperation and courtesies extended to our representatives by the personnel of the Office of the State Comptroller, the Office of Policy and Management, and the state Judicial Department during our review.

Brian DeMilia
Auditor II

Approved:

John C. Geragosian
Auditor of Public Accounts

Robert M. Ward
Auditor of Public Accounts