Death Record Confidence Scoring Tool User Guide

June 2024

Overview

The Death Record Confidence Scoring Tool (DRC Scoring Tool) is a product developed under the Office of Payment Integrity (OPI) to help analytics customers prioritize their research efforts, improve their adjudication processes, and make informed decisions regarding payment eligibility. Using the DRC Scoring Tool, OPI can identify deceased individuals within a dataset more effectively to help customers understand which payments are most at risk of being issued to a deceased payee, and which of the records should be prioritized for their review.

The DRC Scoring Tool analyzes death data available to OPI (see the *Detailed Description* section for a list of current OPI data sources). OPI uses this analysis to determine the accuracy of the death information. The DRC Scoring Tool converts this information into a confidence score based on the likelihood the individual is truly deceased. OPI then uses the confidence score in combination with payment information within the customer's dataset to identify the records that have the greatest potential impact for at-risk dollars, helping customers focus their time on researching and adjudicating records that have the greatest value.

Benefits of the DRC Scoring Tool:

- Customers gain a better understanding of which individuals within their data are deceased based on a confidence score.
- Potentially improper payments to deceased individuals are uncovered and rank-ordered based on potential impact to the agency.
- Customers achieve additional assurance in their payment integrity.
- Customers benefit from less external research and increased efficiency in their adjudication processes.

Features of the DRC Scoring Tool:

The DRC Scoring Tool provides information to customers in a user-friendly file that helps them ingest, prioritize, and act on the insights provided. The file has the following key features:

- Payments and the death record matches are grouped together for each payee, rather than listed as individual line items and mixed in with matches against other payees;
- Interactivity to allow users to quickly navigate through the tool and extract information;
- Research priority of each payee based on the likelihood of death, strength of the match, and dollars at risk;
- Probability of death and the accuracy of the match (death and match confidence);
- Value of at-risk payments; and
- Date a second death source listed the individual as deceased.

Detailed Description

The following walkthrough outlines in detail how the DRC Scoring Tool gathers, processes, and analyzes death data to produce a confidence score for each death record.

1. Death data are gathered from the following data sources:

Data Source	Owner	Description
American InfoSource (AIS) -- Commercial	AIS	Contains obituary and probate information on deceased individuals obtained from over 3,000 funeral homes, thousands of newspapers, and county-level probate records.
Death Master File - Full (DMF-Full)	Social Security Administration (SSA)	Contains records of deaths reported to SSA. The deaths reported to SSA come from many sources, including family members, funeral homes, financial institutions, postal authorities, state information, and other federal agencies.
Department of Defense (DOD) Death Data – Public	DOD	Contains information on active-duty U.S. military and Reserves regarding confirmed or presumed deaths.
Department of State (DOS) Death Data – Public	DOS	Contains records of American Citizens who are deceased or presumed deceased while abroad, as reported by U.S. embassies or consulates upon its receipt of a foreign death certificate or finding of death by a local competent authority.
Payments, Claims, and Enhanced Reconciliations (PACER) Returns and Reclamations Data	Fiscal Service	Contains records of payments made on behalf of the federal government and information on any action taken after a payment has been disbursed; the DRC Scoring Tool uses records of payments that were returned to the agency or reclaimed by the agency due to a death.
Do Not Pay (DNP) Adjudication Data	Do Not Pay	Contains agency adjudication information provided through the DNP Portal; the DRC Scoring Tool uses records of payments that a user from an agency searched in the portal, that returned a match to one of the portal's death data sources (AIS, DMF, DOD, DOS), and that the user adjudicated as improper.
Electronic Verification of Vital Events Fact of Death (EVVE FOD)	The National Association for Public Health Statistics and Information System (NAPHSIS)	EVVE FOD is a premier source of death certificates. EVVE FOD allows users to submit records to be matched against death records from state and jurisdiction owned vital record databases nationwide to inquire if a death certificate exists for those records.

2. Additional factors are used to further validate death and to inform the death record confidence scores for potentially deceased individuals.

3. The data undergo quality checks for the validity of Taxpayer Identification Number (TIN), name, date of birth (DOB), and address. Invalid data may create an inability to match to OPI data sources. TINs, names, dates, and addresses are deemed as valid or invalid according to the following standards:

Element	Invalid	Valid
TIN	Blank, contains text, or does not fit the rules provided by the Social Security Administration (SSA)[1]	Non-blank value, fits SSA rules, [1] and only contains numbers (e.g., "123-45-6789," "123456789")
Name	Blank or contains no letters (e.g., "", "1234")	A sequence of letters (e.g., "John A")
Date of Birth Date of Death	Blank (i.e., "") For EVVE FOD, a date out of range is also invalid (i.e., month value is greater than 12, day value is greater than 31, or year is greater than the current year)	Not blank For EVVE FOD, in-range
Address	Blank (i.e., "")	Not blank

4. Death records associated with a potentially deceased individual are grouped based on how similar the information is in one record compared to the other(s). The similarity score is a number that represents how similar two death records are. The score is based on an algorithm that factors in the reliability of the death data source and the similarity of the TINs, names, dates of birth, and dates of death.

A. TIN similarity is calculated based on how many edits it would take to convert one TIN into the other. Similarity of TINs is weighted higher than similarity of the other key fields. The below table shows how TIN similarity factors into the similarity score between two records.

	Similarity (high to low)	Examples
TIN Similarity	Exact match	"123-45-6789" and "123-45-6789"
	1 digit off	"123-45-6789" and "123-45-6788"

---

[1] The SSA format for SSNs includes the AREA (the first three digits), the GROUP (the middle two digits), and the SERIAL (the final four digits). Valid SSNs: 1) cannot have all zeroes for the AREA, GROUP, or SERIAL; 2) cannot have an AREA be 666; and 3) cannot have the AREA be between 900 and 999.

2 digits off or 1 transposition	"123-45-67**89**" and "123-45-67**77**", "123-45-67**89**" and "123-45-67**98**"
More than 2 digits off or more than 1 transposition	"123-45-6**789**" and "123-45-6**666**", "123-**45**-6789" and "123-**54**-6798"

B. Full name similarity is calculated using algorithms that compare two strings. The more similar the names, the higher the total similarity score.

	Similarity Definition (high to low)	Similarity Score	Examples
Full name similarity	Exact	1	"John Smith" and "John Smith"
	Very similar	0.8 – 1	"John Smith" and "Jon Smith"
	Similar	0.6 – 0.8	"John Smith" and "Jonathan Smyth"
	Different	> 0.6	"John Smith" and "Karen Jones"

C. Date similarity

	Similarity (high to low)	Examples
Date Similarity	Exact	"01-01-2022" and "01-01-2022"
	Similar (less than 35 days apart OR month and day match but year is one character or one transposition off)	"01-**01**-2022" and "01-**31**-2022," "12-31-2021" and "01-01-2022," "01-01-202**2**" and "01-01-202**1**," "01-01-20**20**" and "01-01-20**02**"
	Invalid date	""
	Different	"01-01-2022" and "08-31-2020"

D. In a similar way, middle name, name suffix, ZIP code, and state also impact the similarity score.

E. If the similarity score of a pair of records meets or exceeds the threshold similarity score, they are grouped together. Otherwise, they are considered separate entities.

5. Distinct entities that are similar and as shown above, associated with one TIN are grouped and assigned a unique group ID.

6. Statistics are calculated for each potentially deceased individual which could be associated with either an ungrouped death record or a group of death records. Attributes including the death confidence score, TIN, date of death, and, if found, name, date of birth, and address will appear in the data presented in the Tool.

A confidence score (Low, Moderate, High, or Very High) is created for each group based on the robustness of the data associated with that deceased payee.

7. These data are then matched to OPI death data sources. Any matches are sent to EVVE FOD and returned with the records for which a death certificate was found.

8. After the matching process, the data are imported into an Excel workbook which will become the DRC Scoring Tool deliverable.

Key Columns – Definitions

Payee Tab

Column Name	Definition
Priority	The suggested order that these payees should be reviewed in, based on death and match confidence, associated payment value, and urgency (days until the next anticipated payment disbursal)
Confidence	Based on the estimated likelihood of death (death confidence) and how closely the payments matched to a death record (match confidence); the death confidence is currently based on the number of different death sources the individual is listed on, which particular death sources the individual is listed on, and whether there are any death-related post-payment returns or reclamations listed for this individual; values are either Very High, High, Moderate, or Low
Payee ID	This field is an agency-specific account number that indicates to which account or individual each payment is disbursed; this could be a Master File Payee number
Payee Name	If multiple names are associated with the same payee, this field displays the most frequent or most likely to be correct
SSN	The Social Security number of the deceased; if multiple TINs are associated with the same payee, this field displays the most frequent or most likely to be correct
Death Date	Date of death listed on the death record
Total Pmts	Total number of payments for this payee that were disbursed after the earliest date of death associated with this individual
Total Pmt Amt	The total dollar amount of the payment
Payee Acct Nr	The payment recipient's bank account number
Dual Verification Dt	The earliest date that a second death source listed this individual as deceased
Days Between Pmt and Death	The number of days from the date of death this payment was disbursed
Report Date	The date this current report was run
Death Source List	The death source(s) that provided a match to this record
Payee Status	*(USER INPUT REQUIRED)* – Use the pull-down arrow to select either "Confirmed Alive" or "Confirmed Deceased"
Action	*(USER INPUT REQUIRED)* – Use the pull-down arrow to select either "Halt Payment Series" or "Allow but Monitor"

OPI OFFICE OF PAYMENT INTEGRITY
BUREAU OF THE FISCAL SERVICE

Payments Tab

Column Name	Definition
Payee ID	The Payee Taxpayer Identification Number (TIN); this field is associated with PartyName
Payment ID	The unique payment identifier within the schedule for associating all related records for the payment
ALC	The Agency Location Code for the agency or program making this payment
Payment Type	The Payment Automation Manager (PAM) payment type for this payment, such as Monthly Benefit, Daily Benefit, Vendor, or Miscellaneous
Pmt Amt	The dollar value of this payment
Pmt Date	The date this payment was disbursed
Days Since Death	The number of days between this payment and the most likely death date; in other words, the number of days this individual had been potentially deceased when this payment was disbursed
SSN	The Social Security number of the deceased; if multiple TINs are associated with the same payee, this field displays the most frequent or most likely to be correct
Payee Name	If multiple names are associated with the same payee, this field displays the most frequent or most likely to be correct
Schedule Nr	The agency's schedule number for this payment; determines the batch of payments that are transferred at the same time
ACH Trace Nr	A 15-digit sequence used to identify each ACH payment
Priority	The suggested order that these payees should be reviewed in, based on death and match confidence, associated payment value, and urgency (days until the next anticipated payment disbursal)

Summary Tab

Column Name	Definition
Distinct Screened Payees	The total number of payees that have been screened, regardless of how many payments may appear for each payee
Distinct Matched Payees	The total number of payees that produced matches to the data sources, regardless of how many matches may appear for each payee
Payments Associated with Matched Payees	The number of payments that are associated with the payees that have matches to the data sources
Death Records	The number of death records that matched the list of payees
1-Month Snapshot	The number of payments that are associated with the payees that have matches to the data sources in this reporting time period
Related	The number of payments and dollar amounts that are associated with the payees that have matches to the data sources in previous time periods post date of death
All	The number of payments and dollar amounts that are associated with the payees that have matches to the data sources

Death Record Confidence Scoring Tool Sample Data

The images and detailed descriptions below contain fictitious data.

Upon opening the workbook, you will see four tabs: **Treasury CoverSheet**, **Payee, Payments** and **Monthly Summary**. (Note: if you receive recurring reports, your tabs may differ slightly from those shown below: instead of a Payee tab, you will a see Matches-Current Month tab and a Matches-Recurring tab.)
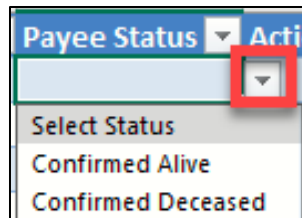
Priority	Confide	Payee ID	Payee Nam	SSN	Death Date	Total Pm	Total Pmt Amt	Payee Acct Nr	Dual Verificatio
1	Very High	1124614023	JAY G CLARK	6663971263	11/23/2023	1	3727.20	6663971263FN	12/20/2023
2	Very High	2488603203	STEPHEN PAL	6403019663	11/24/2023	1	574.74	6403019663FN	12/26/2023
3	Very High	5950413033	JULIUS WISE	2357660623	11/13/2023	1	2718.00	2357660623FN	12/13/2023
4	Very High	1107909043	WILLIAM J HO	8828184483	9/10/2023	1	4002.64	8828184483FN	12/3/2023
5	Very High	2939250003	LUIGI GRAMC	2419314423	11/6/2023	1	2399.39	2419314423FN	12/5/2023
6	Very High	1075909403	DORA C HOW/	4558172943	11/21/2023	1	269.41	4558172943FN	12/20/2023
7	Very High	1379509023	SOREN THYGE	0023414303	11/7/2023	1	5278.34	0023414303FN	12/3/2023
8	Very High	0743510023	PETER CARLEY	2512775523	9/3/2023	1	373.58	2512775523FN	12/3/2023
9	Very High	0262401203	FLOYD KREH	8317307583	12/3/2023	1	1331.67	8317307583FN	
10	Very High	4464713063	JACOB MEYER	2745831523	12/1/2023	1	3466.41	2745831523FN	
11	Very High	9997606003	ED FLOURNO\	2609148523	10/1/2023	1	404.14	2609148523FN	
12	Very High	9954116013	MARY RAASCH	6578948463	11/28/2023	1	465.91	6578948463FN	12/27/2023
13	Very High	8358209203	BLOSSOM M E	4112981543	11/27/2023	1	416.09	4112981543FN	12/20/2023
14	Very High	1542806053	RICHARD YAC	0724187503	12/3/2023	1	2954.00	0724187503FN	12/29/2023
15	Very High	9822516013	DENE RHOAD	2023060523	12/2/2023	1	878.61	2023060523FN	12/27/2023
16	Very High	0047625203	ANN ROACH	8696626983	12/1/2023	1	3056.80	8696626983FN	12/27/2023
17	Very High	1624113213	ALICE OLEARY	0799901703	12/1/2023	1	2024.00	0799901703FN	12/29/2023
18	Very High	7497416013	SARA H ABINA	9960557193	12/1/2023	1	999.00	9960557193FN	12/22/2023
19	Very High	2548209023	MARGARET H	6274198363	11/29/2023	1	3309.20	6274198363FN	12/20/2023
20	Very High	0851603003	RUBE RAWLIN	0592581203	11/27/2023	1	489.77	0592581203FN	12/22/2023
21	Very High	0785216223	ROY CARTER	4859362543	11/3/2023	1	880.32	4859362543FN	12/8/2023

Treasury CoverSheet | **Payee** | Payments | Summary

Payee Tab

The Payee tab lists rows of payments in order of the highest confidence that the payee is deceased to the lowest. Payees with a Confidence Score of Very High, High, or Moderate will be listed first; low confidence matches can be added or removed per request to allow attention to be focused on the most valuable matches.

IMPORTANT! Please use the last two columns to record your adjudication results after conducting your own research, based upon your business rules for payments. Return this spreadsheet to OPI with the results of your research.

    A. The Payee Status column is a user-entered field for each payee row.

Payee Status	Actio
Select Status	
Confirmed Alive	
Confirmed Deceased	

OPI OFFICE OF PAYMENT INTEGRITY
BUREAU OF THE FISCAL SERVICE

When you click in a row of the Payee Status column, a pull-down arrow will appear. Use the arrow to select either **Confirmed Alive** or **Confirmed Deceased** based on your determination.

B. The Action column is a user-entered field for each payee row.
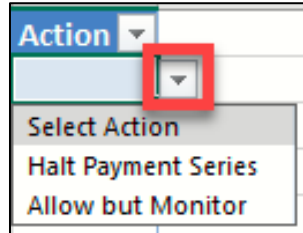
When you click in a row of the Action column, a pull-down arrow will appear. Use the arrow to select either **Halt Payment Series** or **Allow but Monitor** based on your determination.

Payments Tab

The Payments tab lists rows of payments to the payees listed on the Payee tab. Multiple payments to each payee will be listed on successive rows and given the same priority shown on the Payee tab.

Payee ID	Pmt ID	ALC	Record Type	Payment Type	Payment A	Payment Date	Days Since Death	SSN	Full Name
1124614023	110584820786	16149627	Related Paym	MISCELLANEOUS	3727.20	12/29/2023	23	6663971263	JAY G CLARK
2488603203	110515737973	16149627	Related Paym	MISCELLANEOUS	574.74	12/22/2023	15	6403019663	STEPHEN PALUSZKIEWI
5950413033	110584804104	16149627	Match	MISCELLANEOUS	2718.00	12/29/2023	33	2357660623	JULIUS WISE
5950413033	110260341997	16149627	Match	MISCELLANEOUS	2718.00	12/1/2023	5	2357660623	JULIUS WISE
1107909043	109938849916	16149627	Related Paym	MISCELLANEOUS	4002.64	11/3/2023	41	8828184483	WILLIAM J HORNE
1107909043	110584792018	16149627	Match	MISCELLANEOUS	4002.64	12/29/2023	97	8828184483	WILLIAM J HORNE
1107909043	109638686454	16149627	Related Paym	MISCELLANEOUS	4002.64	10/6/2023	13	8828184483	WILLIAM J HORNE
1107909043	110260330240	16149627	Match	MISCELLANEOUS	4002.64	12/1/2023	69	8828184483	WILLIAM J HORNE
2939250003	110260330362	16149627	Match	MISCELLANEOUS	2409.15	12/1/2023	12	2419314423	LUIGI GRAMOLINI
2939250003	110584792137	16149627	Related Paym	MISCELLANEOUS	2399.39	12/29/2023	40	2419314423	LUIGI GRAMOLINI
1075909403	110425449259	16149627	Related Paym	MISCELLANEOUS	269.41	12/15/2023	11	4558172943	DORA C HOWARD
1379509023	110260330416	16149627	Match	MISCELLANEOUS	5278.34	12/1/2023	11	0023414303	SOREN THYGEREN
1379509023	110584792192	16149627	Match	MISCELLANEOUS	5278.34	12/29/2023	39	0023414303	SOREN THYGEREN
0743510023	110584648844	16149627	Related Paym	MISCELLANEOUS	373.58	12/29/2023	104	2512775523	PETER CARLEY
0743510023	109638694385	16149627	Match	MISCELLANEOUS	1494.30	10/6/2023	20	2512775523	PETER CARLEY
0262401203	110584806759	16149627	Related Paym	MISCELLANEOUS	1331.67	12/29/2023	13	8317307583	FLOYD KREH

Treasury CoverSheet | Payee | **Payments** | Summary | ⊕

Summary Tab

The Summary Tab displays a high-level view of the customer data broken out by various categories. See the Summary Tab table in the previous section for definitions of the categories shown.

Summary

Category	Total
Distinct Screened Payees	54,262
Distinct Matched Payees	90
Payments Associated with Matched Payees	195
Death Records	143

Payment Totals

Payment Type	Payments	Total ($)
1-Month Snapshot	91	222,800.20
Related	104	422,456.74
All	195	645,256.94

Payees by Confidence Score

Confidence Text	Total
Very High	56
High	8
Moderate	26

Payments by Confidence Score

Confidence Label	Total Payments (#)	Total Payment ($)
Very High	75	188,638.97
High	92	385,087.85
Moderate	27	68,270.80

Death Sources/Indicators

Death Source/Indicator	Total
AIS Obituary	23
DMF	82
DNP Death Adjudication	1
EVVE FOD	24
PFC	13

OPI OFFICE OF PAYMENT INTEGRITY

BUREAU OF THE FISCAL SERVICE