CalSAWS

Person De-Duplication

Business Process and Communication Protocol

Version 1.1

November 9, 2020

What is a Duplicated Person in relation to CalSAWS Conversion?

During conversion multiple records for a person will be encountered, due to case activity in multiple counties and systems. If a person has applied for or received benefits in more than one county, that person's Client Index Number (CIN) can be associated to multiple person records across the SAWS.   As the three SAWS (LRS, C-IV, and CalWIN) are merged into the single statewide CalSAWS, the conversion process must, where possible, determine which person record is the correct record or "Golden Record" and which person record(s) will be attached, or duplicated, to the Golden Record for historical purposes. (See Appendix A - Key Terms.)

Why is this important?

Approximately 98% of the duplicate person records across the SAWS can be resolved during the conversion process. However, the remaining 2% require manual review. The records that require manual review are person records that are either active in more than one county or active on more than one case within a county. Reviewing these records now will:

- Reduce the number of unresolved duplicates migrated to CalSAWS at conversion.
- Reduce the time counties must spend resolving those duplications post-conversion.
- Reduce the number of MEDS alerts that will be generated post conversion.

CalSAWS Conversion

All person records that meet the CalSAWS Data Retention Policy will convert to CalSAWS regardless of duplication. It is important to reduce the number of duplications prior to conversion. Counties will have to clean up duplications, discrepancies, and alerts in CalSAWS post go-live. However, this process can be time consuming, because counties may have to clean up multiple duplicate records shared with multiple counties while using a process that will involve several steps to complete.

What is the Person De-Duplication Report?

In October 2019, the Conversion Team convened the Person De-Duplication Strategic Mapping Workgroup (User Group) with representatives from all 6 regions to determine how to move forward. The user group decided that the Conversion Team would provide a data cleansing report for the duplicate person records that could not be automated. These reports consist of the 2% duplicate person records mentioned above that will require manual review. Over several months, with the assistance of a dedicated data validation user group, the Conversion Team fine-tuned the report to include tools for counties to sort and filter the data to best fit their needs.

Each county will receive a report that details the active duplicate person records showing for that county. An active duplicate record is defined as a record with a Person Program Status of Active, Approved and/or Pending in more than one case. This can happen to cases within a single county or within two or more counties.

The reason for the duplicates will vary. The most common reason is related to ICT's that are "in-flight". However, there will be instances where an ICT is not the cause. An example might be a college student on Medi-Cal on the parents' case in one county but living in another county with a CalFresh case while away at college. Another example is a customer

receiving CAPI benefits from a CAPI county while receiving other benefits in the county in which they reside. Analysis will be required to determine if a record should be a single, active CalSAWS record or if the duplication is correct.

Data Dictionary

The Conversion Team composed a Data Dictionary for the Person De-Duplication report. The Data Dictionary provides definitions for each of the columns contained in the report. Additionally, the C-IV and CalWIN database source tables and columns are also provided.

Please refer to the Person De-Duplication Report Data Dictionary for details and column definitions in Appendix B.

What can counties do with the report?

The Conversion Team cannot dictate county business process. However, based on the many years of experience from the county staff who have participated in this effort, the counties have identified some "best practices" that have been listed below. The overall goal is to reduce the total number of duplicate person records that are in active status (within a county or in more than one county):

MEDS

- Since de-duplication efforts align with current file clearance efforts, counties might leverage the knowledge and skills of their MEDS Coordinators and MEDS Clerks. File clearance best practices include:
  - Experienced staff inquiring MEDS
  - Experienced staff inquiring SAWS
  - Established procedures for determining correct person record
  - Established procedures for determining correct file clearance
- There are 11 CIN specific alerts generated by MEDS. Counties likely already have business processes in place for reconciling MEDS alerts. These processes can also be leveraged for de-duplication efforts, in conjunction with the de-duplication reports. (See Appendix C- 11 CIN Related MEDS Alerts)
- These efforts can also assist counties to help clear up outstanding MEDS Alerts and may need to be coordinated in order to close open alerts showing in your SAWS system

Inter-County Transfers

- Cases that are marked with a 'Flying ICT Flag' will drop off your list quickly and may not need priority, when ICTs are processed timely. (See Column AF on the Duplicate CIN Report)
- Outstanding ICT's (over 90 days) should be completed as soon as possible, following your county's ICT business processes.
- ICT's in flight should be monitored to make sure that the individual(s) are processed and 'ICT out' records are closed timely. This will help reduce the number of records on the Duplicate CIN Report.
- ICT's received for Medi-Cal and/or CalFresh but then granted under CalWORKS will not disposition the 'ICT out' from the sending county. Contact with the sending county will ensure they close their case, eliminating the duplication.

Data Sorting and Filtering the Person De-duplication Report

- Counties can sort and filter the Person De-duplication Report in several ways to help prioritize the records. Become familiar with the report columns that can be used for sorting and filtering (See Appendix B -

Duplicate CIN Data Dictionary). Some columns that may be of particular interest to counties for sorting or filtering include:

- o Program Pair- column L: can be used to identify and filter programs and/or program pairs in Excel. See example below.
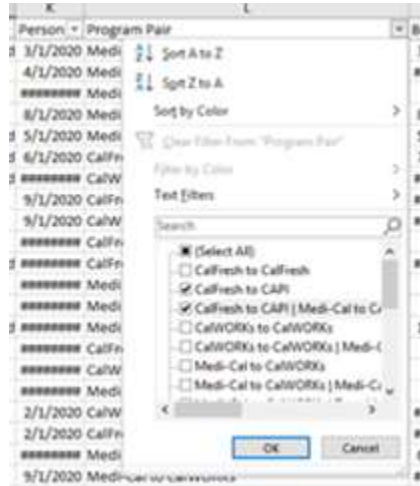
- o Aid Code- Column N: can be used to identify cases that could run in more than one county, i.e.: AAP and Foster Care aid codes.
- o Owning county - column AA: the county with the most recent application date.
- o Pending App >= 90 days - column AC: if the application has been pending 90 days or more.
- o New CIN Record- column AD: highlights the new records included on a report.
- o Flying ICT Flag - column AF: if the incoming/outgoing records for ICT are under 60 days.
- o Active-Active-Active Flag - column AH: when the Disposition Type is showing active in two counties and the Program Status is positive (i.e. Active, Approved, Pending).
- Refer to the Duplicate CIN Data Dictionary in Appendix B for report column descriptions that will help counties use these and other columns for sorting and filtering.

CalWIN Specific Issue

There are records where an individual is marked in CalWIN to "Ignore Eligibility Determination – Yes" on the Collect Case Individual Detail page, Other Information tab. This is typically used for companion cases and for cases with 20+ household members. This indicator may cause CalWIN to not properly determine ineligibility, causing the record to appear as a duplicate.

County Action

Using the above points and suggestions will help counties identify records that need to be closed. It will also highlight those cases that will have to be addressed post-production due to duplications that were unable to be corrected. Counties can then:

- Close erroneous, active or pending records that have a duplicate record with concurrent and correct active or pending Program Status in another county.
- Counties can also work with other counties to request duplications be closed when needed.
- Work this report each month up to your county conversion to CalSAWS.

- Journal/Notate your case/record in your SAWS system. Be sure to observe county business process for Journal/Case Comment entries.

Those cases that cannot be closed will have duplications at conversion and will have to be addressed with a post-conversion clean-up effort.

How will counties communicate during this project?

At the Person De-Duplication Strategic Mapping Workgroup (User Group) session in October 2019, counties requested a process that defines how counties communicate with one another to resolve duplicate person records. At the session, the counties defined the "Owning County" (Column AA) for the Person De-duplication Report. The "Owning County" has the person record with the most recent Application Date for the CIN. The participating counties determined that the "Owning County" is responsible for initial research for a duplicate person record on the report. If the Owning County needs to contact the other county identified on the report, the Owning County will refer to the County Point of Contact list for the Duplicate Person clean-up effort.

De-Duplication Points of Contact List

The Conversion Team has compiled a list of De-Duplication Points of Contact for this effort. The list is available on the CalSAWS SharePoint at: ██████████████████ (You must have access to the Web Portal to access this link.)

Changes to this list should be addressed to Cristy Sharma, at sharmac@calsaws.org.

Personal Identifying Information (PII)

The Person De-Duplication Report was reviewed by the CalSAWS Chief Information Security Officer for compliance with data security guidelines and regulations. While the Person De-Duplication Report does not contain Personal Identifying Information (PII), it does contain MEDS related information. State regulations require CalSAWS and counties passing MEDS related information do so securely and using the minimum information required.

The following are guidelines provided by the Department of Health Care Services (DHCS) that all consortia and counties should follow when transmitting PII[1]:

1. All data transmissions of Medi-Cal PII outside of a secure internal network shall be encrypted using a Federal Information Processing Standard (FIPS) 140-2 certified algorithm that is 128 bit or higher, such as Advanced Encryption Standard (AES) or Transport Layer Security (TLS). It is encouraged, when available and when feasible, that 256-bit encryption be used. *Please note*: all counties that have a current Privacy and Security Agreement (PSA) with DHCS are following the above encryption requirements.
2. Encryption can be end-to-end at the network level, or the data files containing Medi-Cal PII can be encrypted.

This requirement pertains to any type of Medi-Cal PII in motion such as website access, file transfer, and email.

---

[1] Sandra Williams, "2019 Medi-Cal Privacy and Security Agreement (PSA)" ACWDL 19-16: Department of Health Care Services, Medi-Cal Division

Appendix A - Key Terms

Key Terms	
Active/Active Records	Records with a Person Program Status of Active, Approved and/or Pending in more than one county.
Correct CIN	The CIN number identified through MEDS research as the correct CIN for a particular person.
Duplicate Person Records	Multiple person records found in CalWIN, CIV and/or LRS with a single Client Identification Number (CIN) that belong to the same person. This is not to be confused with duplicate records in MEDS where more than one person has been assigned the different CINs.
Golden Record	The Person Record that will be maintained in the CalSAWS system, determined by active status in only one record. Alternatively, it can be selected by worker analysis when there is active status showing in more than one record.
ICT In-Flight	An inter county transfer that has been sent by one county but not yet picked up by the receiving county. (usually pending for up to 60 days)
Person Record	A record in SAWS that belongs to a single individual

Appendix B - Person De-Duplication Report Data Dictionary

Col #	Data Element Name	Data Element Description	CalWIN Source (Table.Column)	C-IV Source / LRS Source (Table.Column)
A	Report County Name	The name of the County for which the report is generated.	N/A	N/A
B	Disposition Type	Based on the program pair combinations and corresponding person program status, a duplicate CIN is associated to one of the following values in this order: Active - Active Active - Pending Pending - Pending	Combination derived from "Person Program Status"	PGM_PERS_DETL.STAT_CODE
C	CIN	Client Index Number associated to the Person.	INDV.CIN	PERS.CIN
D	System	The System (CIV, LRS or CALWIN) to which the county is currently aligned.	Derived - Based on the County Name and currently associated System.	
E	County Name	Name of the County associated to the Person.	Translated to County name using County Code	Translated to County Name using CASE.COUNTY_CODE
F	Case Number	Case Number associated to the Duplicate Person	**AG Programs:** INDV_ASST.CS_ID **Employment Svcs Programs:** ES_STS.CS_ID **No AG or ES records:** CS_INDV_PGM.CS_ID	CASE.SERIAL_NUM_IDENTIF
G	CalWIN Case Status	Status of the Case associated to the Person. Note: this field is only populated for CalWIN Cases as C-IV and LRS do not have a Case level status.	CS.CS_STS_CD	N/A
H	Program	Name of the Program associated to the Person on the Case.	Translated to Program name using: **AG Programs:** INDV_ASST.PGM_TYP_CD **Employment Svcs Programs:** ES_STS.PGM_TYP_CD **No AG or ES records:** CS_INDV_PGM.PGM_TYP_CD	Translated to Program name using PGM.PGM_CODE
I	Program Status	Status of the Program.	Translated to Status name using CS_PGM.PGM_STS_CD	Translated to Status name using PGM_DETL.STAT_CODE
J	Person Program Status	Status of the Person on the Program.	Translated to Status name using: **AG Programs ('AP', 'DC'):** AG_INDV_ELIG_RSLT.ELIG_STS_CD **AG Programs ('PE', 'CN', 'WD', 'DE'):** CS_PGM.PGM_STS_CD	Translated to Status name using PGM_PERS_DETL.STAT_CODE

Col #	Data Element Name	Data Element Description	CalWIN Source (Table.Column)	C-IV Source / LRS Source (Table.Column)
			Employment Svcs Programs: ES_STS.STS_CD	
K	Person Program Status Date	Effective date of the status of the Person on the Program.	**AG Programs:** CS_INDV_PGM.STS_DT **Employment Svcs Programs:** ES_STS.STS_DT **No AG or ES records:** CS_PGM.STS_DT	PGM_PERS_DETL.BEG_DATE
L	Program Pair	Represents the type of program pair identified between duplicate persons records. Note: The Program Type described below represents the name of the Program and will display within this field on the report. Same Program Type to Same Program Type 'Medi-Cal to CalWORKs' 'Medi-Cal to CAPI' 'CalFresh to CAPI' 'Transitional CalFresh to ' Any other Program Type	Derived - based on Program Type	
M	Begin Date of Aid	Date from which the Person was receiving aid in the Program (Latest Date).	**AG Programs:** INDV_ASST.AID_BGN_DT **Employment Svcs Programs:** ES_STS.STS_DT **No AG or ES records:** CS_INDV_PGM.STS_DT	PERS_APP.AID_BEG_DATE
N	Aid Code	Aid Code for the Program in which the Person was receiving aid (Latest Aid Code).	**AG Programs:** INDV_ASST.AID_CD **Employment Svcs Programs:** N/A	AID_CODE.AID_CODE (for Non-MC and Non-FC Aid Codes) BUDGET_PERS.AID_CODE (for MC Aid Codes) FC_AID_CODE.AID_CODE (for FC Aid Codes)
O	App Date	Application Date of the Person associated to the Program.	CS_PGM.APP_ID; AR_SAWS1_FORM.APP_DT	PERS_APP.APP_DATE
P	Role	Role of the Person within the Program.	Translated to Role name using: **AG Programs:** AG_INDV_ELIG_RSLT.PRM_PRTCP _STS_CD	Translated to Role name using PGM_PERS_DETL.ROLE_CODE
Q	In the Home	Represents if the person is In the Home within the case or not (Y/N)	CS_INDV.IN_HOM_SW	hh_stat.STAT_CODE (value of 'IH' = 'Y' else 'N')

Col #	Data Element Name	Data Element Description	CalWIN Source (Table.Column)	C-IV Source / LRS Source (Table.Column)
R	CalHEERS Case Number	Associated CalHEERS Case Number, if applicable.	CS_INDV_CH_XREF.CH_CS_ID	CH_CASE_INFO.CH_CASE_NUM_IDENTIF
S	Worker ID	Worker Number assigned to the Program.	CS_PGM.CSLD_ID, SE_CSLD_PRM_USR.USR_ID	STAFF_WRKR.WRKR_NUM_IDENTIF
T	Office	Office to which the Worker is assigned.	SE_USR_UNT_OFF.USR_ID, SE_OFF.OFF_ID	OFFICE.OFFICE_NAME
U	Unit	Unit in which the office is present.	SE_USR_UNT_OFF.USR_ID, SE_UNT.UNT_ID	UNIT.UNIT_NAME
V	District	District in which the office is present.	N/A	Translated to District name using OFFICE.DISTR_CODE
W	ICT Incoming Date	Date on which the Incoming ICT was received, if it was an ICT case.	CS_PGM.APP_ID; AR_SAWS1_FORM.ENT_DT	ICT.BATCH_DATE
X	ICT Outgoing Date	Date on which the ICT was sent out, if it was an ICT case.	IN_ICT_STS_STS_OUT.TRGR_DT	ICT.BATCH_DATE
Y	Confidentiality Type	Confidentiality Type, if the case was marked as Confidential.	CS_SPL_SITN.SITN_TYP_CD, RT_SPL_CHAR.SPL_CHAR_CD	Translated to Confidentiality Type name using CONFID_DETL.TYPE_CODE
Z	Within/ Outside County dup	Represents if the CIN record is a duplicate within the same county or a duplicate with a different county.	Derived - If the duplicate records exist within the same County then 'Within County Duplicate', else 'Outside County Duplicate'.	
AA	Owning County	County that will be taking lead to address the duplicate CIN records.	Derived - The County which has the most recent Application Date for the CIN. This column will show dual ownership when the Application Date in both counties are the same. Example: Alameda | San Jose	
AB	Number of days App Pending	For Pending Programs, the number of days between the Application Date and the date the report was generated.	Derived - Days between Application Date and date the report was generated.	
AC	Pending App >= 90 days	Switch ('Y') set for the applications that are pending for more than 90 days.	Derived - If Number of days App Pending >= 90, then 'Y'	
AD	New CIN Record	Represents if the CIN is newly added to the current report since the previous report (Y/N).	Derived - If the CIN, on the current report, did not exist within the previous report, then 'Y'. Note: For the baseline report, this field is blank for all records.	
AE	App Age Flag	Flag set to "1" for all records that match the CIN with application pending over 90 days (see column "Pending App >= 90 days	Derived - Identifies all matching CIN records associated with CIN record(s) that have a Y in the column "App Pending >= 90" thereby displaying the record(s) with matching CIN even if only one record has 'Y' value. This flag's purpose is to hide/eliminate records (cases) for which the VALUE TO WORK RATIO IS VERY LOW	

Col #	Data Element Name	Data Element Description	CalWIN Source (Table.Column)	C-IV Source / LRS Source (Table.Column)
AF	Flying ICT Flag	Flag set to "1" for all records that match the CIN where the ICT date is "in-flight" meaning the Incoming or Outgoing date is less than or equal to 60 days which is a window described to give Counties time to address backlogs of ICTs. Inside this window counties' normal processes are assumed to address these individuals/cases without the need for the extra effort associated with this Duplicate CIN report/process		Derived - Identifies all matching CIN records associated with CIN record(s) where the date in the columns "ICT Incoming Date" or "ICT Outgoing Date" is less than or equal to 60 days old (in-flight ICT). This column supports displaying/hiding the record(s) with matching CIN even if only one record has an "in-flight" ICT date. The purpose of this flag is to hide/eliminate records (cases) for which the VALUE TO WORK RATIO IS VERY LOW
AG	Active Active Negative Flag	Flag is set to "1" for all records that match the CIN where the "Disposition Type" is 'Active-Active' and the "Program Status" is a negative status (i.e. Discontinued, Withdrawn, Canceled, etc.)		Derived - Identifies all matching CIN records associated with CIN record(s) where the "Disposition Type" is 'Active-Active' and the "Program Status" is negative. This column supports displaying the record(s) with matching CIN even if only one record meets the criteria and for which the VALUE TO WORK RATIO IS VERY HIGH
AH	Active Active Active Flag	Flag is set to "1" for all records that match the CIN where the "Disposition Type" is 'Active-Active' and the "Program Status" is a positive/pending status (i.e. Active, Approved, Pending)		Derived - Identifies all matching CIN records associated with CIN record(s) where the "Disposition Type" is 'Active-Active' and the "Program Status" is positive. This column supports displaying the record(s) with matching CIN even if only one record meets the criteria and for which the VALUE TO WORK RATIO IS VERY HIGH

Appendix C - 11 CIN Related MEDS Alerts

Alert#	Alert Title	Alert Description
505	CIN Based County ID Contains Erroneous Serial Number	No description or action specified by MEDS
1087	Trans has CIN but no Check-digit	No description or action specified by MEDS
1088	CIN/CIN Check Digit Conflict	No description or action specified by MEDS
1503	CIN/MEDS-ID Conflict	This alert is generated when: *　there is a CIN on the transaction that is known to MEDS; *　the MEDS-ID reported on the transaction does not match the MEDS-ID associated 　　with that CIN on the MEDS CIN Cross Reference File; and *　the MEDS-ID reported on the transaction is not a previously used MEDS-ID 　　associated with the CIN Cross Reference File MEDS-ID.
1504	CIN/MEDS ID Versus County ID/MEDS ID Conflict	This alert is generated when: ·　there is a CIN on the transaction that is known to MEDS; ·　the MEDS-ID reported on the transaction matches the MEDS-ID associated with 　　that CIN on the MEDS CIN Cross Reference File; ·　the County ID on the transaction is known to MEDS; and ·　the MEDS-ID reported on the transaction does not match the current MEDS-ID 　　associated with the County ID on the MEDS County ID Cross Reference File.
1508	SCI CIN/MEDS ID Conflict- MEDS ID Change Req'd	The CIN was NOT submitted on the transaction; the transaction MEDS ID was used to find the CIN on the Statewide Client Index. A CIN found on the Statewide Client Index, but that CIN is not associated with the MEDS ID on the transaction.
1509	Two MEDS Records Have Same CIN- Chaining Req'd	Two different MEDS IDs on MEDS are using the same Client Index Number (CIN). These two MEDS records are potentially duplicate records.
1514	CIN Not known to MEDS and not Primary CIN on SCI	No description or action specified by MEDS
2163	Add Trans Does Not Contain Client Number- CIN005ERROR	No description or action specified by MEDS
4824	SSN Verification Not Posted Due to CIN/MEDS ID Conflict	This message may be issued along with message 4801, 4802, 4803, 4804, 4805, 4806, 4807, 4808 or 4809. It indicates that the SSN verification information for this SSN will not update MEDS because the SSN points to a different CIN than it did when the SSN verification request was initiated.
9028	Potential Duplicate CIN- Must Be Resolved	No description or action specified by MEDS