|  |
| --- |
|  |

|  |
| --- |
| **Updating the Hepatitis B Registry - Standard Operating Procedures** |
|  |
|  |
|  |
|  |
| **California Department of Public Health (CDPH)** |
|  |
|  |

Table of Contents

[I. Purpose 1](#_Toc389070968)

[II. CDPH Contacts 1](#_Toc389070969)

[III. Overview 2](#_Toc389070970)

[Data Dictionaries 2](#_Toc389070971)

[Code Repositories in Github 2](#_Toc389070972)

[Match Procedures 3](#_Toc389070973)

[IV. Standard Operating Procedures 3](#_Toc389070974)

[Step 1: Update CalREDIE Dataset 4](#_Toc389070975)

[Step 2: Update Non CalREDIE Dataset 5](#_Toc389070976)

[Step 3: Update PHPP Dataset 6](#_Toc389070977)

[Step 4: Update CMRs Dataset (Optional) 7](#_Toc389070978)

[Step 5: Combine CalREDIE Dataset, Non CalREDIE Dataset, and PHPP dataset into a Morb file. 7](#_Toc389070979)

[Step 6: Update SF eFTP Dataset 9](#_Toc389070980)

[Step 7: Update Registry SAS Code to Reflect Newly Updated Datasets 9](#_Toc389070981)

[Step 8: Run Registry SAS Code (Part 1) 10](#_Toc389070982)

[Step 9: Review the scores in the “trough” in SAS Graph output 10](#_Toc389070983)

[Step 10: Run Registry SAS Code (Part 2) 11](#_Toc389070984)

[V. References 12](#_Toc389070985)

[VI. Appendix A. Steps to Prep and Combine Source Data to Make Main01 and Main02 files 14](#_Toc389070986)

[VII. Appendix B. Definition of Registry SAS Programs 15](#_Toc389070987)

[VIII. Appendix C. Definition of Permanent Registry Datasets 16](#_Toc389070988)

# 

# Purpose

The purpose of this document is to provide direction and assistance to California Department of Public Health (CDPH) staff who conduct viral hepatitis surveillance. This document will describe the procedures for updating the viral hepatitis B registries on an annual basis.

# CDPH Contacts

The following individuals at CDPH play key roles in viral hepatitis surveillance:

STD Control Branch

* Heidi Bauer, STD Control Branch, Chief
* Michael Samuel, STD Control Branch, Epidemiology and Surveillance Section, Chief
* Joan Chow, STD Control Branch, Epidemiology Unit, Chief
* Rachel McLean, STD Control Branch, Office of Adult Viral Hepatitis Prevention, Chief

Immunization Branch

* Kathleen Harriman, Immunization Branch, Vaccine Preventable Disease Epidemiology Section, Chief
* Erin Murray, Immunization Branch, Epidemiologist Supervisor

# Overview

In 2012, CDPH created hepatitis B case registries that the STD Control Branch uses to characterize the epidemiology of viral hepatitis in California. The hepatitis B case registries require a deduplication and matching methodology described in detail in the publication by Glenn Wright, *“Probabilistic Record Linkage in SAS”*[*1*](#_ENREF_1).

## Data Dictionaries

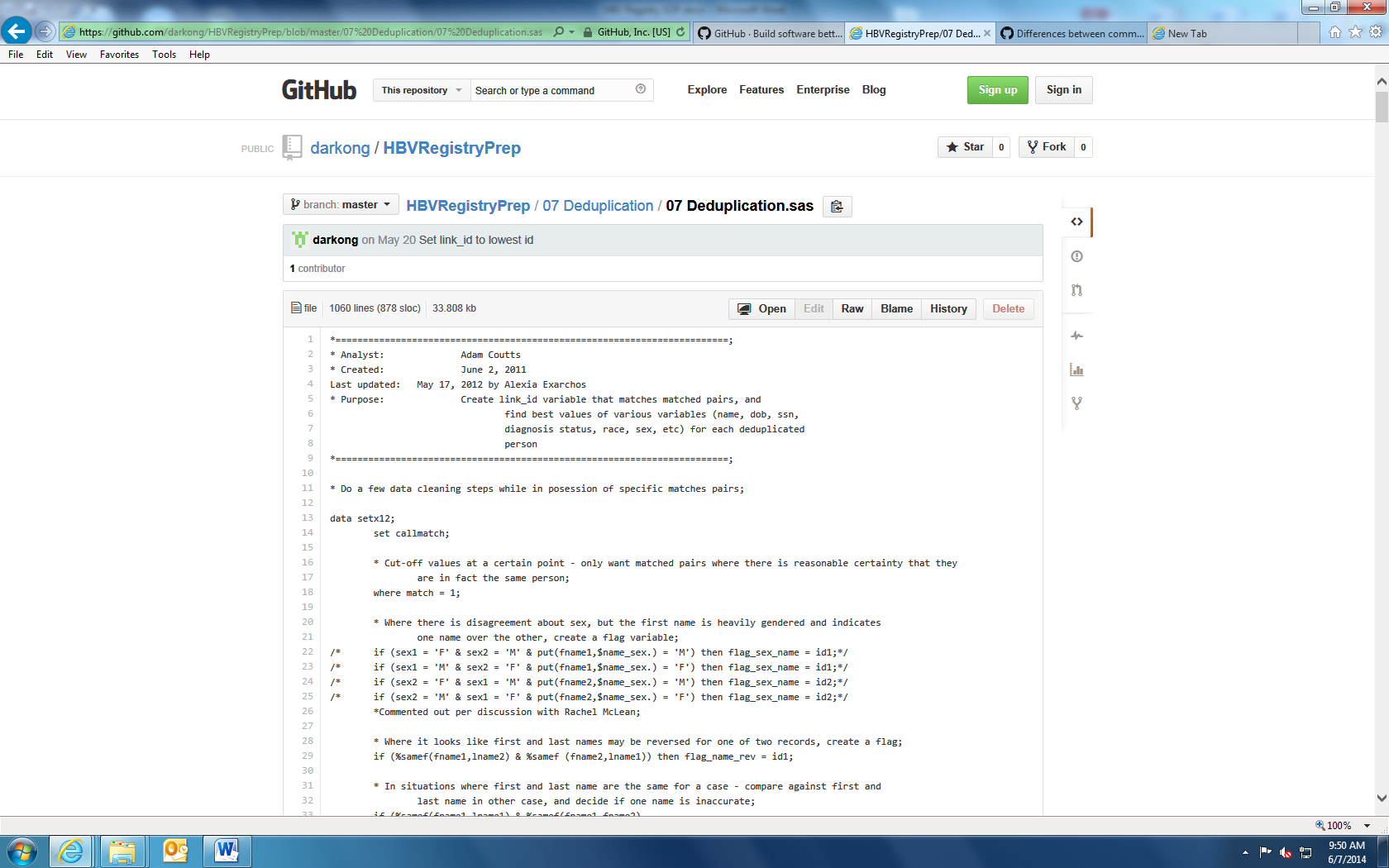
There are two hepatitis B registry that contains both chronic cases since 1988 and acute cases since 2010. Data dictionaries for each of the registries can be found here (*R:\State Surveillance\HBV Registry\HBVRegistryPrep*); variables in the registries will not be explained in detail in this document.

## Code Repositories in Github

The Registry Code (programs 01A-08), all the source data preparation codes used in steps 1-5, the data dictionary, and this SOP are contained in the folder *R:\State Surveillance\HBV Registry\HBVRegistryPrep.* This folder is connected to GitHub which means that any changes made to the contents can be linked to the cloud and you can see past changes and the documented rationale for those changes. You can review the past changes to this folder at the following website:

<https://github.com/darkong/HBVRegistryPrep>

To check the history of the files, click the file you are interested in, and click history:



The site will then show you the dates when the code was changed, will highlight the changes made, and display comments regarding the reason for the changes.

## Match Procedures

1. In order to update the HBV registries correctly, you must first compile the raw data in **Steps 1-6**.
2. Once you have completed **Step 6**, close your SAS session to clear out the work datasets and macros. Then run **Steps 7-11.**

# Standard Operating Procedures

**Step 1:** Update CalREDIE Dataset

**Step 2:** Update Non CalREDIE Dataset

**Step 3:** Update PHPP Dataset

**Step 4:** Update CMRs Dataset (Optional)

**Step 5:** Combine CalREDIE Dataset, Non CalREDIE Dataset, PHPP Dataset, and CMRs Dataset into a Morb file

**Step 6:** Update San Francisco (SF) electronic File Transfer Platform (eFTP) dataset

**Step 7:** Update registry SAS code to reflect new data

**Step 8:** Run registry SAS code (Part 1)

**Step 9:** Review the scores in the “trough” in SAS Graph output

**Step 10:** Run registry SAS code (Part 2)

**Step 11:** Create permanent registry datasets

Note: Unlike the HCV registry, many of the source files are compiled into a Morb file before the registry codes officially start. Steps 1-4 require cleaning of source data for records that were created Jan 1 of the current year to the present. Step 5 will create a year specific morb file with records from Jan 1 of the current year to the present as well as a new cumulative morb file with all Non San Francisco data for all years. Thr new cumulative morb file is created by combining the current year-specific data with the old cumulative morb file so that the records with a create date of Jan 1 of the current year or later will only come from the current year-specific morb file (see Appendix A). The new cumulative morb file along with the cumulative San Francisco data will be fed into the registry program to create the main01 and main02 files.

Also, because the source datasets are recreated and merged into a morb file before the official registry code (programs 01A to 08), the Cartesian matching in program 06 will be performed across all records old and new. This means that link\_ids may change from year to year in the HBV registry. In contrast, the HCV registry keeps the new records separate from the old and therefore allows Cartesian matching to be restricted so that old data is not matched again to old data. This means that link\_ids in the HCV registry will have more consistency from year to year.

Lastly, one other major difference between the HCV and HBV registries is that the HBV registry includes acute cases from 2010 to the present. The HCV registry only includes chronic cases. Acute and chronic HBV reports can be identified through specific indicator variables (see dictionary).

## Step 1: Update CalREDIE Dataset

***Input:***

Download Case Report Exports using the create date range of Jan 1 of the newly added year to the current date for the following UDFs below. HEPC info is also brought in, but currently it is not used in the HCV registry code.

HEPBCLICR

HEPBLABCR

HEPCCLICR

HEPCLABCR

Save the files in the following folders:

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\CalREDIE Cumulative\yyyy\acHBVCLI.mdb

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\CalREDIE Cumulative\yyyy\acHBVLAB.mdb

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\CalREDIE Cumulative\yyyy\acHCVCLI.mdb

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\CalREDIE Cumulative\yyyy\acHCVLAB.mdb

Also, download from the Data Distribution Portal the system tab exports for all Disease Grouping Incidents using the start date of Jan 1 of the new year to the current date. This will give the system tab info for all IZB diseases, but only the hepatitis cases will be used. The file will download as a tsv file, but convert it to an Excel file and save it as:

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\CalREDIE Cumulative\yyyy\sysextract.xlsx

***Output:***

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\CalREDIE Cumulative\yyyy\calrediecumul.sas7bdat

Note: yyyy = four digit year

***Run:***

R:\State Surveillance\HBV Registry\HBVRegistryPrep\PrepSourcedata\PrepCalREDIEdata.txt

1. When running the file, be sure to change the highlighted directory to reflect the new data:

libname calredie "R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\CalREDIE Cumulative\2012";

Note: yyyy = four digit year

## Step 2: Update Non CalREDIE Dataset

***Input*:**

Contact Tamara Henessey-Burt to obtain the hepatitis data feeds from jurisdictions not on CalREDIE (other than San Francisco) with create dates from Jan 1 of the new year to July 2013 of the following year. Previously she has given SAS files for the AVSS counties, Los Angeles, and San Diego. Save those files here:

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\NonCalREDIE Cumulative\yyyy\hepalameda.sas7bdat

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\NonCalREDIE Cumulative\yyyy\hepavss.sas7bdat

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\NonCalREDIE Cumulative\yyyy\hepavssb.sas7bdat

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\NonCalREDIE Cumulative\yyyy\hepla.sas7bdat

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\NonCalREDIE Cumulative\yyyy\heplab.sas7bdat

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\NonCalREDIE Cumulative\yyyy\hepsd.sas7bdat

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\NonCalREDIE Cumulative\yyyy\hepsdb.sas7bdat

***Output:***

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\NonCalREDIE Cumulative\yyyy\noncalrediecumul.sas7bdat

***Run:***

R:\State Surveillance\HBV Registry\HBVRegistryPrep\PrepSourcedata\PrepNonCalREDIEdata.txt

Note: yyyy = four digit year

Counties that were not on CalREDIE previously may have transitioned to CalREDIE since the last update

Because of upcoming changes to the AVSS data structure, it is likely that this code will need to some intensive rewriting to ensure the correct variables and formats are used. Please consult with Tamara Henessy-Burt to ensure code is complete.

## Step 3: Update PHPP Dataset

***Input*:**

Maternal data collected for the perinatal hepatitis B prevention program (PHPP) from the beginning to the present should be included in our HBV registry. Contact the perinatal hepatitis B prevention program (PHPP) coordinator to confirm that the most up-to-date PHPP data and is located here:

Z:\PerinatalHepB\Datafiles\ PHPP APRIL 2011\_be\_B4ON2005\_ACTIVE.accdb

Z:\PerinatalHepB\Datafiles\ PHPP\_PAST5\_BACKEND\_ACTIVE.mdb

Run the following program:

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\PHPP Cumulative\yyyy\combine – All Moms Enrolled.sas

To create the following dataset, which will be the basis for the PHPP dataset:

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\PHPP Cumulative\yyyy\allmomsenrolled.sas7bdat

***Output:***

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\PHPP Cumulative\yyyy\phppcumul.sas7bdat

***Run:***

R:\State Surveillance\HBV Registry\HBVRegistryPrep\PrepSourcedata\PrepPHPPdata.txt

Note: yyyy = four digit year

The structure of the PHPP database may have changed since the last update. Check with Billy Luong of IZB if you have any questions about variables.

## Step 4: Update CMRs Dataset (Optional)

***Input*:**

Add any data from loose CMRs to the Excel file below. However, there should be no further CMRs to enter. Therefore, if no additional Hepatitis B CMRs have been added there is no need to run this step

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\CMRs\yyyy\Hepatitis B CMRs.xlsx

***Output:***

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\CMRs\yyyy\cmrs.sas7bdat

***Run:***

R:\State Surveillance\HBV Registry\HBVRegistryPrep\PrepSourcedata\PrepCMRdata.txt

Note: yyyy = four digit year

## Step 5: Combine CalREDIE Dataset, Non CalREDIE Dataset, and PHPP dataset into a Morb file.

***Input*:**

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\CalREDIE Cumulative\yyyy\calrediecumul.sas7bdat

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\NonCalREDIE Cumulative\yyyy\noncalrediecumul.sas7bdat

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\PHPP Cumulative\yyyy\phppcumul.sas7bdat

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\CMRs\yyyy\cmrs.sas7bdat

***Output:***

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\Year Specific Morbs\yyyy\hepbmmddyyyy

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\morbyyyy

Note: mm = two digit month; dd = two digit day; yyyy = four digit year

If there is no additional CMR data added to the cmrs file since 2012, then there is no need to create a new file and you can reference the old cmr file at:

R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\CMRs\yyyy\cmrs.sas7bdat

***Run:***

R:\State Surveillance\HBV Registry\HBVRegistryPrep\PrepSourcedata\createmorbyyyy

1. Change the highlighted text to set the old morb dataset to be the last morbfile. For example, to update the registry with 2013 data, the old morbfile should be set to morb.morb2012

**data** oldmorb;

set morb.morb (drop = resolution);

**run**;

1. The data step that creates the newmorb file restricts based on create date. Change the highlighted create dates to reflect Jan 1 of the current year as the start date and the current date as the end date:

**data** newmorb;

set calredie

noncalredie

phpp

cmrs;

…

if prxmatch("m/mexican/oi", race) > **0** then race= 'White';

if report in ('PHPP', 'CMRS') or (report in ('Morb', 'CalREDIE') and dtcreate > mdy(**12**,**31**,**2011**) and dtcreate < mdy(**10**,**1**,**2013**));

**run**;

1. Change the highlighted names of the output files of the last two data steps (creating year-specific morb and new morb files) to reflect current dates and years. Also change the highlighted date of the last data step to be Jan 1 of the current year so that the only data for Jan 1 of the current year comes from the current year-specific morb.

**data** hbvdest.hepb11222013;

set newmorb;

if dis in ('HEP-B', 'HEP-B-CR');

**run**;

**data** morb.morb2012;

set hbvdest.hepb11222013

oldmorb (in = a);

if a = **1** and dat > mdy(**12**,**31**,**2011**) then delete;

**run**;

## Step 6: Update SF eFTP Dataset

***Input:***

Contact Amy Nishimura at SFDPH to get the most recent SF Hepatitis data: <https://eft.dhcs.ca.gov/EFTClient/Account/Login.htm>

***Output:*** R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\SF\yyyy\cdph\_sf\_chronicb.sas7bdat

***Run:***

1. If you do not already have access, request permission to the SF eFTP site from Rachel McLean ([Rachel.McLean@cdph.ca.gov](mailto:Rachel.McLean@cdph.ca.gov)) or Kathleen Harriman at [Kathleen.Harriman@cdph.ca.gov](mailto:Kathleen.Harriman@cdph.ca.gov), depending on who your supervisor is.
2. Log onto the SF eFTP site (<https://eft.dhcs.ca.gov/EFTClient/Account/Login.htm>) using your username and password.
3. Download the SF SAS Dataset cdph\_sf\_chronicb.sas7bdat to the following location: R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\SF\yyyy\, using the following naming convention: cdph\_sf\_chronicb.sas7bdat, where yyyy = four digit year

Note: yyyy = four digit year

## Step 7: Update Registry SAS Code to Reflect Newly Updated Datasets

***File Location:*** R:\State Surveillance\HBV Registry\HBVRegistryPrep\01\_Standard\_Header\01\_Standard\_Header.sas

***Run:***

1. Update the latest iteration of the input SAS program to import the most recent downloaded data
2. Change the highlighted text to reflect the updated SAS datasets:

**data** MorbFile;

set "R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\Morb2012.sas7bdat";

**run**;

**data** SFFTP;

set "R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\SF\2012\cdph\_sf\_cases\_chronicb.sas7bdat";

**run**;

1. Save the SAS program with an updated date in the same file location

## Step 8: Run Registry SAS Code (Part 1)

***File Locations:***

1. R:\State Surveillance\HBV Registry\HBVRegistryPrep\01\_Standard\_Header\01\_Standard\_Header.sas
2. R:\State Surveillance\HBV Registry\HBVRegistryPrep\01A Macros\01AMacros.sas
3. R:\State Surveillance\HBV Registry\HBVRegistryPrep\03 Datasource Preparation\03 Datasource Preparation.sas
4. R:\State Surveillance\HBV Registry\HBVRegistryPrep\03 Datasource Preparation \04 Prison Macro\04 Prison Macro.sas
5. R:\State Surveillance\HBV Registry\HBVRegistryPrep\05 Merging\05 Merging.sas
6. R:\State Surveillance\HBV Registry\HBVRegistryPrep\06 Matched Pairs\06 Matched Pairs.sas

***Run:***

Run the most current iteration of the registry SAS code in the following order:

1. 01 Standard Header
2. 01A Macros
3. 03 Datasource Preparation
4. 04 Prison Macro
5. 05 Merging
6. 06 Matched Pairs --- see **Step 5** below for details on reviewing the output ---

**Note:** This code should be run in sections to ensure that there aren’t any error messages or warnings that could result in erroneous data.

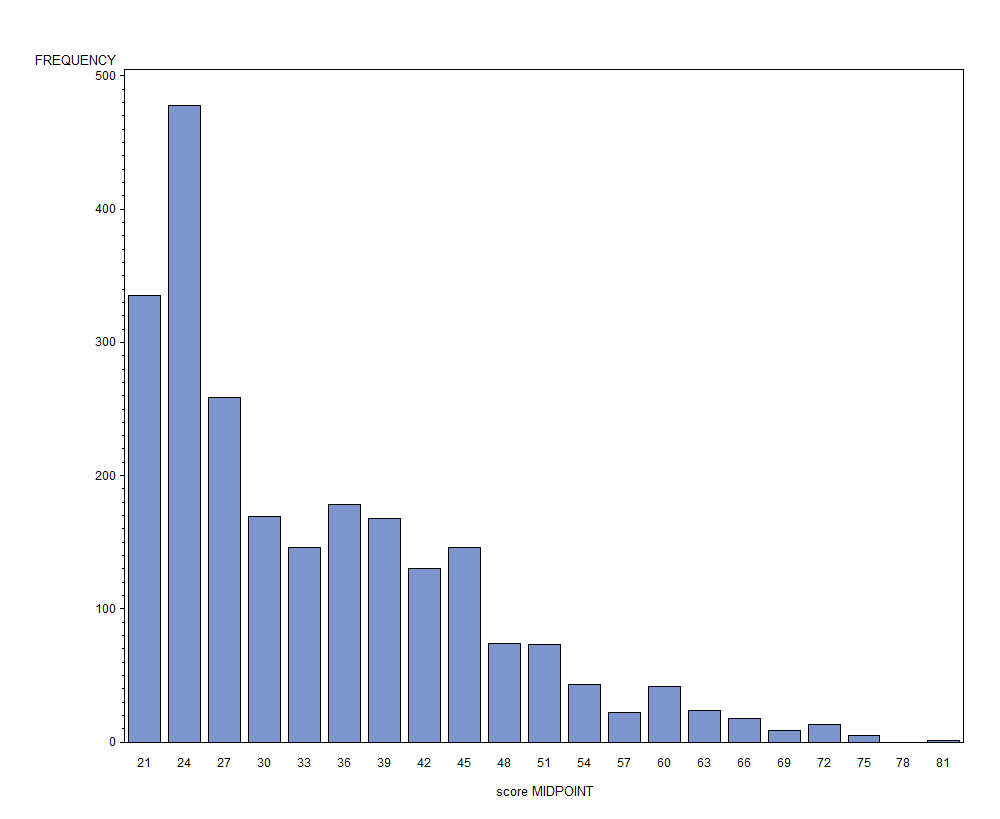
## Step 9: Review the scores in the “trough” in SAS Graph output

***File Location:*** R:\State Surveillance\HBV Registry\HBVRegistryPrep\06 Matched Pairs\06 Matched Pairs.sas

***Run:***

1. In program 06 Matched Pairs, SAS Graph will produce a distribution curve of the scores from the linked pairs dataset (Chart 1)
2. Identify the scores that reflect the dip before the bell curve (trough)
3. Confirm that the scores that contain the trough are close to 28.5 – 33. The scores do not have to be identical to the trough, but should be roughly the same.
4. If the trough scores are dramatically different, this is most likely due to error in the newly uploaded datasets. The Epidemiologist will need to review the updated data to look for data entry errors, make corrections, and start the process over.

**Chart 1. Printout of SAS Graph is Program, “06 Matched Pairs”**



The bars to the right of the trough represent cases that matched

The arrows indicate the scores of the “trough” before the rise where cases are deemed a match. These scores identify the range of scores with potential matches that are further analyzed in the SAS program

The bars to the left of the trough represent cases that did not match

## Step 10: Run Registry SAS Code (Part 2)

***File Locations:***

1. R:\State Surveillance\HBV Registry\HBVRegistryPrep\07 Deduplication\07 Deduplication.sas
2. R:\State Surveillance\HBV Registry\HBVRegistryPrep\08 Final Steps\08 Final Steps.sas

***Run:***

1. Run the most current iteration of the registry SAS code in the following order:
   1. 07 Deduplication
   2. 08 Final Steps
2. Save and close the programs in the same locations using the same file names

**Note:** This code should be run in sections to ensure that there aren’t any error messages or warnings that could result in erroneous data.

# References

1. Wright G. Probabilistic Record Linkage in SAS. *SAS Users Group*. San Francisco: Western Users of SAS Software (WUSS); 2011.

A copy of the paper can be found here: *G:\State Surveillance\Surveillance Analyses References*

# Appendix A. Steps to Prep and Combine Source Data to Make Main01 and Main02 files

Year-Specific Morb

CMRs\*

PHPP

Non CalREDIE

CalREDIE

New Morb

Main01

Old Morb

(before current year)

Main02

SF Data

= datasets with multiple years of data

= datasets with 1 year’s worth of data

Dataset locations\*\*:

|  |  |
| --- | --- |
| CalREDIE | R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\CalREDIE Cumulative\yyyy |
| Non CalREDIE | R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\NonCalREDIE Cumulative\yyyy |
| PHPP | R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\PHPP Cumulative\yyyy |
| CMRs | R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\CMRs\yyyy |
| Year-Specific Morb | R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb\Year Specific Morbs\XXXX |
| Old Morb | R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morb[yyyy -1].sas7bdat |
| New Morb | R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\Morbyyyy.sas7bdat |
| SF Data | R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\Source Data\SF\ yyyy |
| Main01 | R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\main01\_chronichbv.sas7bdat |
| Main02 | R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets\main02\_chronichbv.sas7bdat |

\* CMRs were a one-time addition for the update with 2012 data and is not expected to be included any longer.  
\*\*yyyy = current year, yyyy-1 = prior year.

# Appendix B. Definition of Registry SAS Programs

* **01 Standard\_Header:** This program assigns the directories and runs all of the formats that will be used in later programs.
* **01A Macros:** This program installs macros used in later programs.
* **03 Datasource Preparation:** This program revises and standardizes the code to merge all of the source datasets together (morbfile and SF eFTP).
* **04 Prison Macro:** This program calls in the prison macro which determines whether a given case was reported from a state prison.
* **05 Merging:** This program merges source datasets into one main dataset, processes the data, and recodes variables in preparation for the merge.
* **06 Matched Pairs:** This program creates matched pairs and “Linked Pairs” and “Look Carefully” SAS datasets stored here: *R:\State Surveillance\HBV Data Analysis\HBV Match 2012\Datasets* of cases to deduplicate in the next program. This program applies scores to potential matched variables; Lastname, Firstname, Date of Birth, and Social Security Number.
* **07 Deduplication:** This program creates a “link\_id” variable that links matched pairs and finds the best value for the following variables: name, date of birth, date of death, social security number, diagnosis status, race (if available), sex, prison, and location for each individual case. This methodology is described in more detail in the paper, *“Probabilistic Record Linkage in SAS”*[*1*](#_ENREF_1).
* **08 Final Steps:** This program breaks the final, best-choice-demographic dataset called, “setx15” from the previous program, into two final datasets (person-level and event-level).
* **09 Registry Creation:** This program creates a master HBV registry of chronic cases with their acute diagnosis and report dates and a master HBV registry of acute cases with their chronic diagnosis and report dates.

# Appendix C. Definition of Permanent Registry Datasets

The permanent registries are as follows:

* **Main01\_chronichbv:** This is the permanent chronic HBV Registry and is a SAS dataset of cumulative chronic HBV deduplicated person-level data combined with acute data since 2010. Each record represents an individual reported with chronic or acute HBV and each individual is represented once in the dataset. There are variables to indicate whether person had acute HBV records, chronic HBV records, or both.
* **Main02\_chronichbv:** This is the permanent chronic HBV Registry and is a SAS dataset of cumulative chronic HBV and acute HBV (since 2010) event-level data. Each record represents a report of chronic or acute HBV. Multiple records corresponding to the same individual indicate that one individual was reported with chronic HBV multiple times. The Main02\_chronichbv records can be linked to the Main01\_chronichbv records via the link\_id variable.