

# ARKANSAS APCD DATA USERS GROUP

**Kenley Money**  
**Director of Information Systems**  
**Architecture**

07.28.2021



**H**EALTHCARE  
**T**RANSPARENCY  
**I**NIITIATIVE

# Agenda

- Welcome
- Topics
  - Methodologies for Managing High Collision Hash IDs
  - Methodologies for Developing Project Specific Events/Episodes
  - Data Field Coverages
  - New Medicaid Data Changes
  - Latest APCD Release Information and Data Tips
- Questions/Discussion

# Arkansas APCD Team

- Kenley Money, MA, MFA — Director of Information Systems Architecture/ACPD Director, ACHI
- Nichole Sanders, PhD — Assistant Director of Analytics, ACHI
- Mike Motley, MPH — Director of Analytics, ACHI

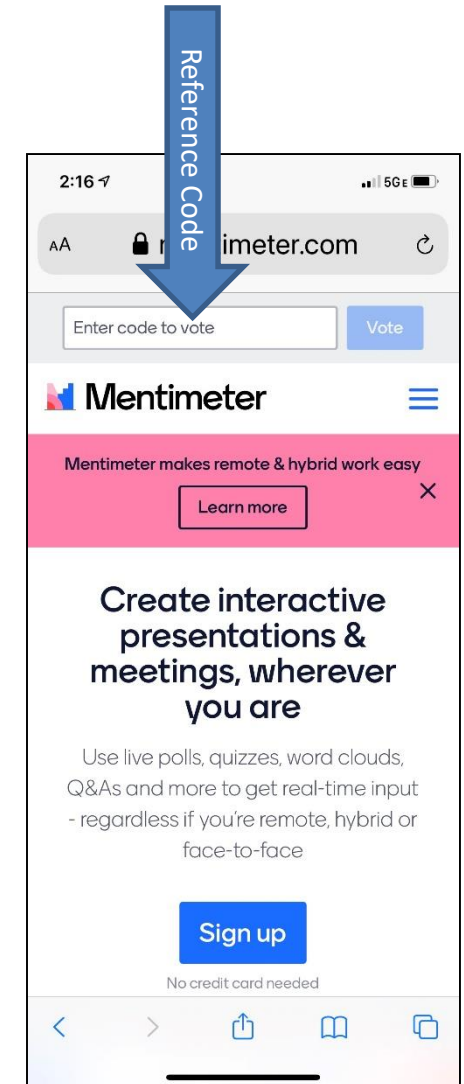
# Something new

- Collecting feedback using Mentimeter, an interactive presentation software
- Slides are placed throughout the presentation with questions about the topic.
- A reference number is at the top of the feedback slide.

Go to [www.menti.com](https://www.menti.com) and use the code 8703 7025

# Mentimeter

- Go to [Mentimeter.com](https://Mentimeter.com) on your phone or [Menti.com](https://Menti.com) on a different browser.
- When prompted, enter the reference code found at the top of the feedback slide in the box here.
- You will see a screen with the question and answer options.
- Enter your feedback.



# Methodologies for Managing High Collision Hash IDs

- An APCD Unique ID\* – or Hash ID – is a ‘hashed’ value that represents the concatenated member last name and date of birth
- Combining the Hash ID with gender can be used as a proxy to identify a unique individual – when supported by other information
- The purpose of this approach is to determine a close approximation of the number of distinct individuals in a study population using a combination of gender and Hash ID

\*Data Element ID for APCD Unique ID: ME998.

# Methodologies for Managing High Collision Hash IDs

- Hash IDs are not as precise as using a complete set of personally identifiable information
- ‘Collisions’ occur when last names, dates of birth, and genders are the same for more than one person
- Collision influencers
  - Twins (approximately 1–3% rate)
  - Common last names over time
  - Common dates of birth within groups or years
  - Data quality errors

# Methodologies for Managing High Collision Hash IDs

- Collisions, while present, do not occur often and have a very small impact on the overall data set
- Collision rate in a randomly created analytic set is expected to be around 3.5% and potentially lower for a smaller sized set

*Collision rates can vary depending on the size and design of the analytic dataset.*



# Methodologies for Managing High Collision Hash IDs

- Some Hash IDs have large frequency counts

Top 10 Hash IDs (No Submitting Entity Association)	Gender	Count
1yDhYf2ILOh98nS+ZtWv4tZWckdaOwqZf3xnnnVLNVE=	F	4,936
1yDhYf2ILOh98nS+ZtWv4tZWckdaOwqZf3xnnnVLNVE=	M	2,838
BRhv0qu1IFibrrKW1fwdzzZkj9uf8rhes+SKDaHivyA=	U	298
86LA0+at6RSsFOecByqJQEG+s9VFOpabcsqIUmj3XGk=	U	225
oxdvja0i6yMkZ9r2HEzqlhuC31ngwQob0+/qwdDjrQA=	U	317
qlXRdk60oQwYpuSK0zAdlg487utqlJJRAyHv8dfchDM=	M	191
MfC1CF3AWQoX+5eeNFEoHOMRvju5AMneU50zJbF7Tr8=	U	179
/jM3tzC8pGgF85Y3808LeEqhkuQQMYQHPOzDoY8kTmw=	M	156
U6M9RNokWVeY1ysdTZgh56k/oJty33FmbGfg92BczeM=	U	154
u7wonVC7GrIU2NIXu82Codh1FbTQ9RaYR/KlhD9cOOc=	U	152

Can these be correct?  
Would an enrollee really have that many member records?

Are these really collisions?

\*These counts represent the number of active or latest member enrollment (versionrank 1) records.

# Methodologies for Managing High Collision Hash IDs

- Hash IDs / submitting entity groups with high row counts

Submitting Entity	Hash ID	Gender	Row Count
99EBD1	1yDhYf2ILOh98nS+ZtWv4tZWckdaOwqZf3xnnnVLNVE=	F	4,936
99EBD1	1yDhYf2ILOh98nS+ZtWv4tZWckdaOwqZf3xnnnVLNVE=	M	2,838
99HSM1	U6M9RNokWVeY1ysdTZgh56k/oJty33FmbGfg92BczeM=	M	3
99CAR1	U6M9RNokWVeY1ysdTZgh56k/oJty33FmbGfg92BczeM=	M	56
99CAR1	U6M9RNokWVeY1ysdTZgh56k/oJty33FmbGfg92BczeM=	U	154
99CAR1	U6M9RNokWVeY1ysdTZgh56k/oJty33FmbGfg92BczeM=	F	7
60054A	7u7P27dJ4ehf8rKexncDZ3qQhio9+6NQABaM3wfpB9U=	M	105
47155	7u7P27dJ4ehf8rKexncDZ3qQhio9+6NQABaM3wfpB9U=	M	1
99MCD1	ivMwwCBwDFZUd5pe3nU/czwvMgWghPnxFzJq4o8v6bl=	F	55
83470	ivMwwCBwDFZUd5pe3nU/czwvMgWghPnxFzJq4o8v6bl=	F	14

\*These counts represent the number of active or latest member enrollment (versionrank 1) records.

# Methodologies for Managing High Collision Hash IDs

- Reasons for high count Hash IDs include but are not limited to:
  - Data quality issues
    - Blank names
    - Default names
    - Default dates
    - Common names
  - Treatment for severe illnesses
    - Home health/skilled nursing facilities
    - Mental health facilities, including residential facilities
    - Pharmaceutical treatment for chronic conditions
- While some Hash IDs occur in high quantities, they are not always collisions

# Methodologies for Managing High Collision Hash IDs

- Overall, the number of unique Hash IDs/Gender combinations associated with multiple Member IDs is relatively low.

~2M Hash IDs  
have between 2  
to 5 member IDs

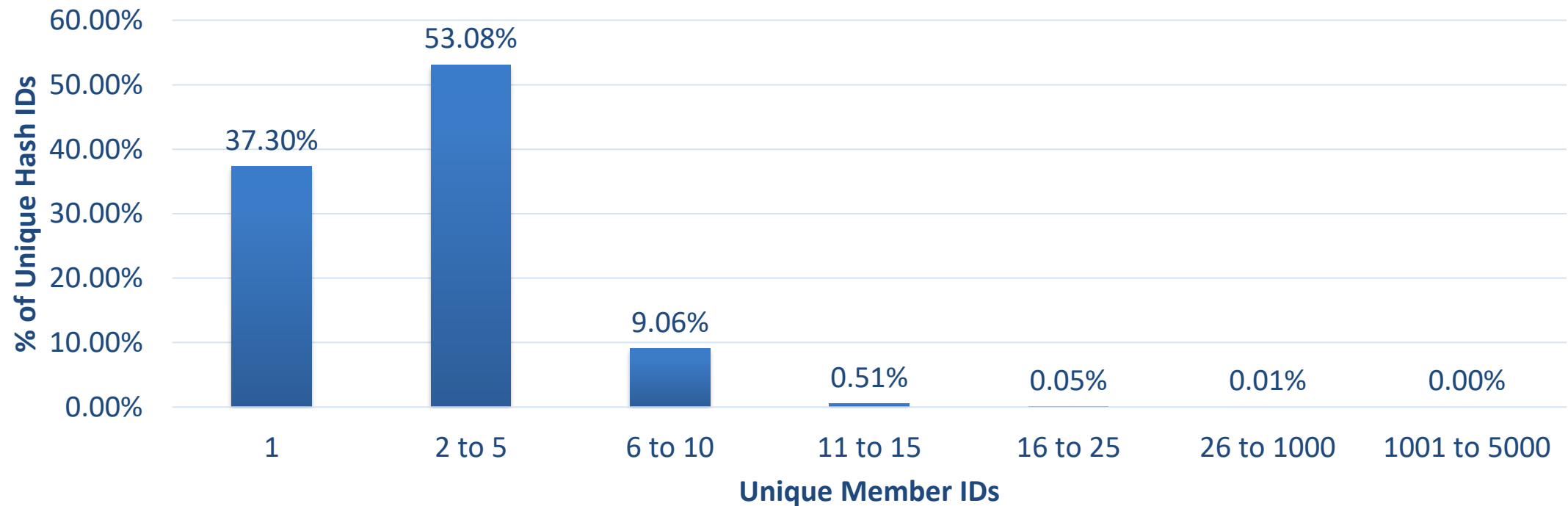
Number of Member IDs	# of Hash IDs*
1	1,465,055
2 to 5	2,085,039
6 to 10	355,829
11 to 15	19,852
16 to 25	1,976
26 to 1000	234
1001 to 5000	2
Total Unique Hash IDs	3,927,987

★ Member IDs are the concatenated Submitting Entity ID (ME001) and Member Number (ME107)

\*These counts represent the number of active or latest member enrollment (versionrank 1) records.

# Methodologies for Managing High Collision Hash IDs

- Distribution of Hash IDs to Member IDs – All APCD



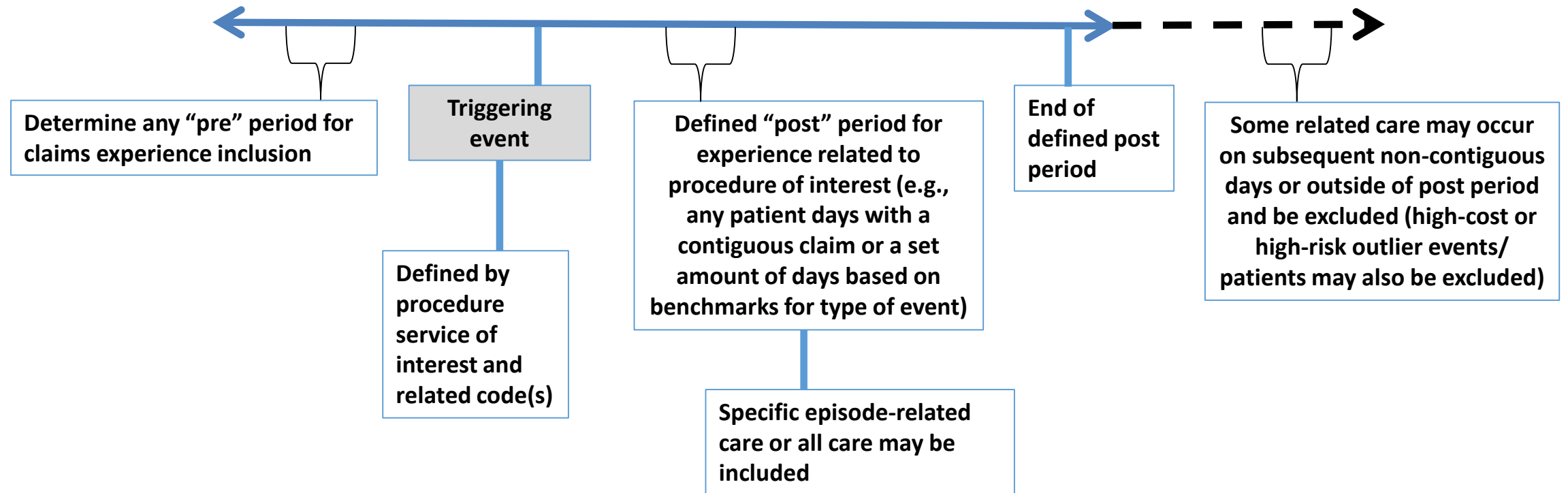
# Methodologies for Managing High Collision Hash IDs

- Differences in Member IDs are most often because a member was enrolled with different carriers over time
- Member IDs can also change because of name changes, plan changes, moves, etc.
- Having multiple member IDs associated with a single Hash ID is not always a problem

# Methodologies for Managing High Collision Hash IDs

- Collision reduction can be managed in several ways:
  - Analytic datasets focused on rare or specialized conditions are typically small and often don't include collisions; collision reduction steps not usually required
  - Records associated with collisions within large analytic datasets focused on common conditions could be omitted and still maintain a viable analytic dataset
  - If collisions are problematic in an analytic dataset, applying a collision reduction methodology is recommended
- April 2020 Arkansas APCD Data Users Group Deck has additional collision management strategies: [www.arkansasapcd.net/Docs/1392/](http://www.arkansasapcd.net/Docs/1392/)

# Methodologies for Developing Project Specific Events/Episodes





# Methodologies for Developing Project-Specific Events/Episodes – Examples

- Emergency department visits
  - Leading to inpatient stays
- Colonoscopies
- Tonsillectomies
- Deliveries

# Data Field Coverages

- Data field coverages for each file type can be found on the Arkansas APCD [website](#)
- The [Arkansas APCD Data Element Frequency Counts](#) provide row counts by data element value for all submitting entities combined
- Click on each file type to download an Excel file with data element counts

# New Medicaid Data Changes

- Data issues recently identified in Arkansas Medicaid Data:
  - Facility Type – MC037
    - This field had previously been hard coded as “02” on all records for all years
    - Data was updated with correct values in the 19B release (January 2020), however new records received in later submissions continued to be coded with “02” in error
    - This field has been corrected for all years of Arkansas Medicaid medical claims
    - NOTE: As of 2017, the value “02” represents telemedicine; the correction data will still contain the value “02” for telemedicine



# New Medicaid Data Changes

- Data issues recently identified in Arkansas Medicaid Data:
  - Service Provider Data
    - The original logic supporting service provider data fields preserved NULL values when the service provider information was not available at the claim line level
    - The logic has been revised to select service provider data from the claim header level when the claim line level service provider data is not available
    - This change will result in more service provider level data for use in analytics
    - These fields have been corrected for all years of Arkansas Medicaid medical claims



# New Medicaid Data Changes

Service Provider Data Elements	
MC024 -SRVC_PRVDR_NUMB	MC031 - SRVC_PRVDR_SUFFIX
MC025 -SRVC_PRVDR_EIN	MC032 - SRVC_PRVDR_SPECIALTY
MC026 - SRVC_NPI	MC033 - SRVC_PRVDR_CITY
MC027 - SRVC_PRVDR_ENTITY_TYPE	MC034 - SRVC_PRVDR_STATE
MC028 - SRVC_PRVDR_FIRST_NAME	MC035 - SRVC_PRVDR_ZIP
MC029 - SRVC_PRVDR_MIDDLE_NAME	MC070 - SRVC_PRVDR_COUNTRY_CODE
MC030 - SRVC_PRVDR_LAST_NAME	MC108 - SRVC_PRVDR_STREET



# New Medicaid Data Changes

- Data issues recently identified in Arkansas Medicaid Data:
  - Ethnicity – ME025
    - The original logic applied a general categorization to ethnicity by categorizing all values into “33” (Not Hispanic or Latino – Other or Blank (no race selected)) or “34” (Hispanic or Latino – Other or Blank (no race selected))
    - Arkansas Medicaid has updated the population of the ethnicity flag with more specific data that aligns with the values described in [Appendix I – Ethnicity](#)
    - Ethnicity fields have been corrected for all years of Arkansas Medicaid member data



# Release Information

- Available APCD data
  - Current APCD Data: Jan. 1, 2013, through June 30, 2020
  - Next build in progress
    - Includes data Jan. 1, 2013, through December 31, 2020
    - Estimated completion: September–October, 2021



**Always check the Arkansas APCD  
Data Issues and Tips page for the  
latest information!**



# Data Tips

- Utilize searchable Arkansas APCD [data dictionaries](#) & [tip sheets](#)
- Highlights (be sure to review them all!):
  - Resolved issues
  - Updated Tips/Issues
  - Featured Tips/Issues
    - [Issue 0080: Delta Dental Open Enrollment Segments](#)
    - [Issue 0083: EBD Pharmacy Provider NPI Mapping](#)
    - [Issue 0082: Aetna versioning](#)
    - [Issue 0087: Inpatient and Institutional Definition Expansion](#)

# APCD Technical Support

- Reach out to [adrs@achiapcd.atlassian.net](mailto:adrs@achiapcd.atlassian.net) for questions about data requests, data use, or pricing
- Something special!
  - If you are interested in a one-on-one meeting with the Arkansas APCD team, reach out to us through our technical support email above

# Call to Action

- Sign up for ACHI Newsletter
- Follow on social media: ACHI and the Arkansas Healthcare Transparency Initiative featuring the Arkansas APCD



- Check out the blog posts on ACHI website
- Next users group meeting: October 27, 2021

**SUBSCRIBE  
FOR UPDATES**

**[achi.net/newsletter](http://achi.net/newsletter)**

