TB & COVID-19 Matching
November, 2021

Evan Timme | AZ-TB Epi
Disclosure
-no conflicts
Surveillance system matching [not EMR]

\[ \text{TB} \text{ [confirmed & counted in AZ]} \]

\[ \iff \]

\[ \text{COVID-19} \text{ [confirmed/probable in AZ]} \]

Flexible matching – no MPI

*MPI - master Pt index
TB
[~2/100,000]

COVID-19
[>17,000/100,000]

Matched
[>20%]

*matched via SAS program
**only captures AZ TB and AZ COVID-19

Matching

UID
or
COMPGED function
or
SPEDIS function
or
FIND function
or
General Surveillance ID
Matching

<table>
<thead>
<tr>
<th>UID</th>
<th>Name/DOB/Sex combo</th>
</tr>
</thead>
<tbody>
<tr>
<td>or</td>
<td>YASIIN BEY</td>
</tr>
<tr>
<td>or</td>
<td>12/11/1973</td>
</tr>
<tr>
<td>or</td>
<td>Male [m=1, f=2]</td>
</tr>
<tr>
<td>or</td>
<td>YSBY1211731</td>
</tr>
<tr>
<td>or</td>
<td>*missing/unk =&gt; ‘9’</td>
</tr>
</tbody>
</table>

or

SPEDIS function
or
FIND function
or
General Surveillance ID
Matching

**UID**
or
**COMPGED function**
or
**SPEDIS function**
or
**FIND function**
or
**General Surveillance ID**

Name/DOB/Sex combo

YSBY1211731 = YSBY1211731

TB **UID** = COVID **UID**

**Exact match**

96% match
Matching

- UID
- or
- COMPGED function
- or
- SPEDIS function
- or
- FIND function
- or
- General Surveillance ID
generalization of Levenshtein edit distance

Measures dissimilatory

*higher #, greater diff

First name ≤200, AND
Last name ≤200, AND
DOB ± 1-day
Matching

UID
or
COMPGED function
or
SPEDIS function
or
FIND function
or
General Surveillance ID

TB
YASIN BEY
12/11/1973

COVID
YASIN BAY
12/11/1973

MATCH
F_name = 20
L_name = 100

MATCH
F_name = 40
L_name = 40

91% match
Matching distance query

UID
or
COMPGED function
or
SPEDIS function
or
FIND function
or
General Surveillance ID

Computes Asymmetric spelling
*higher #, greater diff

First name $\leq 40$, AND
Last name $\leq 40$, AND
DOB $\pm$ 1-day
Matching

UID or COMPGED function or SPEDIS function or FIND function or General Surveillance ID

**MATCH**
- F_name = 8
- L_name = 33

**NO MATCH**
- F_name = 33
- L_name = 66

<table>
<thead>
<tr>
<th>TB</th>
<th>COVID</th>
</tr>
</thead>
<tbody>
<tr>
<td>YASIN</td>
<td>YASIN</td>
</tr>
<tr>
<td>BEY</td>
<td>BEY</td>
</tr>
</tbody>
</table>

95% match
Matching

UID
or
COMPGED function
or
SPEDIS function
or
FIND function
or
General Surveillance ID

TB name contained in COVID-19 name [vv]

\[
((TB \ F\_name \ in \ COVID \ F\_name) \\
\text{or} \\
(COVID \ F\_name \ in \ TB \ F\_name)) \\
\text{AND} \\
((TB \ L\_name \ in \ COVID \ L\_name) \\
\text{or} \\
(COVID \ L\_name \ in \ TB \ L\_name)) \\
\text{AND} \\
\text{DOB} \pm 1\text{-day}
\]
Matching

UID
or
COMPGED function
or
SPEDIS function
or
FIND function
or
General Surveillance ID

98% match

TB
YASIIIN
BEY
12/11/1973

COVID
YASVIN
BAY
12/11/1973

NO MATCH
F_name ➔ NO
L_name ➔ NO

MATCH
F_name ➔ YES
L_name ➔ YES
Matching

UID
or
COMPGED function
or
SPEDIS function
or
FIND function
or
General Surveillance ID

Surveillance system ID

Parent ID across 2+ morbidities
*requires manual process

75% match
Matching

UID --------------------------------- 96%
or
COMPGED function -------------- 91%
or
SPEDIS function ------------- 95%
or
FIND function ------------- 98%
or
General Surveillance ID ----- 75%
THANK YOU

Evan Timme  |  TB Epi
evan.timme@azdhs.gov  |  480-244-6643

Coded in SAS 9.4