

1. An updated SAS macro to differentiate ICD-9-CM and ICD-10-CM records
2. Use of dual coded data to assess comparability ratios for ICD-based surveillance/morbidity measures

Eduard Poltavskiy, MS
Patrick Romano, MD MPH
NAHDO/UC Davis team
February 3, 2016

Background

- The Health Insurance Portability & Accountability Act of 1996 (HIPAA) named certain types of organizations as covered entities, including health plans, clearinghouses, and most providers.
- In HIPAA regulations, the Secretary of HHS adopted certain standard transactions for Electronic Data Interchange (EDI): claims and encounter information, payment and remittance advice, claims status, eligibility, enrollment and disenrollment, referrals and authorizations, coordination of benefits and premium payment.
- If a covered entity conducts one of the adopted transactions electronically, then it must adhere to the content and format requirements for that type of transaction.
- HHS modified the standard code sets for EDI procedures by adopting ICD-10-CM for diagnosis coding and ICD-10-PCS for inpatient hospital procedure coding, effective October 1, 2015 for all covered entities.

A SAS macro to differentiate ICD-9-CM and ICD-10-CM records

Rationale

- Some data reporting entities (e.g., property and casualty insurers, disability, workers compensation, employee health clinics, correctional health clinics) are not covered by HIPAA and did not switch to ICD-10-CM/PCS on 10/1/2015.
- Many (but not all) payers required claims with dates of service spanning the 10/1/2015 implementation date to be split so that the services prior to 10/1/2015 were billed separately using ICD-9-CM codes; services on and after 10/1/2015 were billed separately using utilize ICD-10-PCS codes.
- Some claims processors report relatively high percentages of records with ≥ 1 invalid or incomplete ICD-10-CM/PCS code.
- Invalid or incomplete codes may not be recognized in morbidity or surveillance measure definitions based on specific ICD-10-CM/PCS codes

A SAS macro to differentiate ICD-9-CM and ICD-10-CM records

Assumptions

- A rapid method is needed to verify that a record was coded using the expected code set, and to reject noncompliant records.
- Invalid or incomplete codes may need to be rejected or revised (e.g., adding 7th character)
- An individual record will include either ICD-9 or ICD-10 codes, but never (validly) both.
- E codes are not valid principal diagnoses in ICD-9-CM; V codes are not valid principal diagnoses in ICD-10-CM.

A SAS macro to differentiate ICD-9-CM and ICD-10-CM records

ICD-9-CM (v32) included 14,567 diagnosis codes and 3,882 procedure codes, whereas ICD-10-CM/PCS (FY 2016) lists 69,823 diagnosis codes and 71,974 procedure codes. Coding structures differ somewhat (Table 1) and around 500 diagnosis codes are duplicated in the both code sets.

Table 1. Key Differences between ICD-9-CM and ICD-10-CM/PCS Code Structures*

| | ICD-9-CM | ICD-10-CM |
|---------------------------------|--|--|
| Diagnosis Code Structure | 3-5 characters Character 1 is numeric or alpha (E** or V) Characters 2-5 are numeric | 3-7 characters Character 1 is alpha Character 2 is numeric Characters 3-7 can be alpha or numeric |
| | ICD-9-CM Procedure Codes | ICD-10-PCS |
| Procedure Code Structure | Minimum of 3 digits, maximum of 4 digits. | Seven alphanumeric characters. Each character contains up to 34 possible values. |

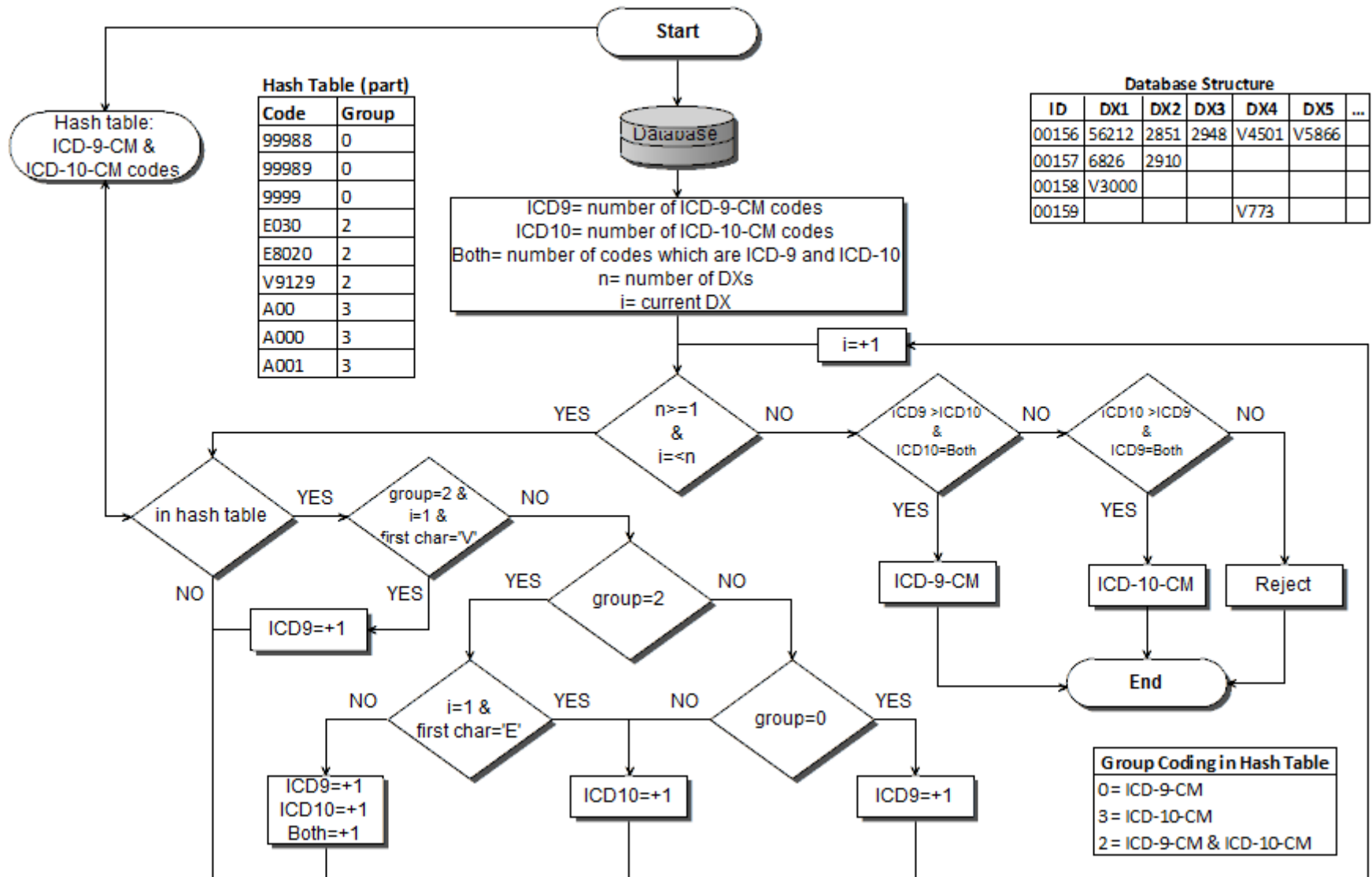
* http://www.cdc.gov/nchs/icd/icd10cm_pcs_background.htm

** 'E' code cannot be a principal Dx in ICD-9-CM.

ICD2016_CLASS_DxPr.zip package

1. **%ICD_CLASS** software: a SAS macro to differentiate ICD-9-CM and ICD-10-CM diagnosis records.
2. **%ICD_PCS_CLASS** software (**recently added**): a SAS macro to differentiate records with ICD-9-CM procedure codes and ICD-10-PCS procedure codes.
3. User Guide.

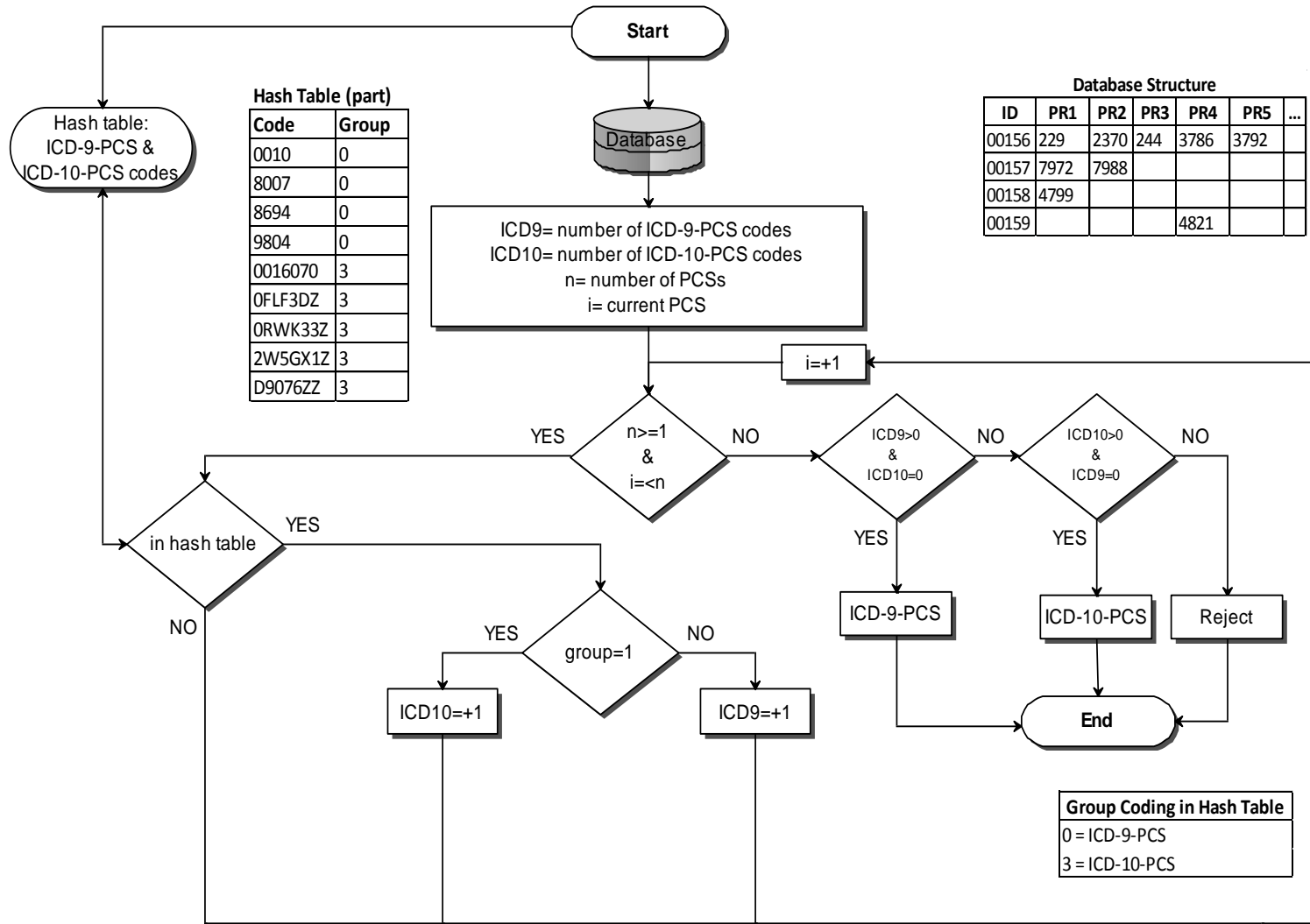
Algorithm to Classify a Record with Diagnosis Codes



The algorithm counts number of ICD-9-CM, ICD-10-CM, and codes that are both ICD-9-CM and ICD-10-CM for each record. Then, it compares those counts to classify a record. A record will be flagged as unclassified (Rejected) if

- 1) it has no valid codes;
- 2) the number of valid ICD-9-CM codes equals to the number of valid ICD-10-CM codes;
- 3) the number of codes found in the both coding system (Both) is different from the number of codes classified as ICD-9-CM or ICD-10-CM.

Algorithm to Classify Records with Procedure Codes



The algorithm counts the number of valid ICD-9-CM procedure codes, valid ICD-10-PCS, and invalid codes that could not be found in either ICD-9-CM or ICD-10-PCS code sets. Then, it compares those counts to classify a record.

A record will be flagged as Unclassified (Rejected) if

- 1) it has no valid procedure codes or;
- 2) it has a valid ICD-9-CM procedure code(s) and a valid ICD-10-PCS code(s).

Artificial Data, MIX Dataset, n=17

| | | | | | | | Four new column will be added to the dataset | | | |
|----|---------|---------|---------|---------|---------|---------|---|--------------------------|-------|--------|
| ID | PR1 | PR2 | PR3 | PR4 | PR5 | PR6 | Number of invalid codes | Number of valid codes | ICD_9 | ICD_10 |
| 1 | 3521 | 3525 | 3961 | 3562 | 3959 | 3931 | 0 | 6 | 1 | 0 |
| 2 | 3582 | 3885 | 3961 | ERROR | | | 1 | 3 | 1 | 0 |
| 3 | 3582 | 3885 | 3561 | 3961 | | | 0 | 4 | 1 | 0 |
| 4 | 3582 | 5A1221Z | 3961 | 3790 | | | 0 | 4 | 0 | 0 |
| 5 | 3563 | 3961 | 3572 | | | | 0 | 3 | 1 | 0 |
| 6 | 3572 | | | | | | 0 | 1 | 1 | 0 |
| 7 | 05B60ZZ | B24DYZZ | 3582 | 3961 | | | 0 | 4 | 0 | 0 |
| 8 | 02RF07Z | 02RH0KZ | 02U507Z | 021K0KQ | 02QM0ZZ | 5A1221Z | 0 | 6 | 0 | 1 |
| 9 | 02170ZP | 5A1221Z | 02LR02T | 02U507Z | 07BM0ZZ | 02U70JZ | 1 | 5 | 0 | 1 |
| 10 | 02170ZP | 02160ZP | 02U507Z | 02LR02T | E08809 | 02QP0ZZ | 1 | 5 | 0 | 1 |
| 11 | 3961 | 05L00ZZ | 05L10ZZ | 05L10ZZ | 5A1221Z | 02160ZP | 0 | 6 | 0 | 0 |
| 12 | 02QJ0ZZ | 06540ZZ | 06500ZZ | 02QG0ZZ | 5A1221Z | | 0 | 5 | 0 | 1 |
| 13 | | | 02QL0ZZ | 02SR0ZZ | | | 0 | 2 | 0 | 1 |
| 14 | | | 02Q40ZZ | 5A1221Z | | | 0 | 2 | 0 | 1 |
| 15 | | | | 02Q40ZZ | 02LR02T | | 0 | 2 | 0 | 1 |
| 16 | | | 02Q40ZZ | 5A1221Z | B24DZZ4 | 0DJ08ZZ | 0 | 4 | 0 | 1 |
| 17 | | | | | 02QM0ZZ | 5A1221Z | 0 | 2 | 0 | 1 |

| | | | | | |
|-------------|--------------|-------------|--------------------|-------------|----------------------|
| CODE | Invalid code | CODE | Code from ICD-9-CM | CODE | Code from ICD-10-PCS |
|-------------|--------------|-------------|--------------------|-------------|----------------------|

Definition of the Classification

| ICD_9 | ICD_10 | Definition |
|-------|--------|-------------------------------------|
| 1 | 0 | record was classified as ICD-9-CM |
| 0 | 1 | record was classified as ICD-10-PCS |
| 0 | 0 | record was not classified |

```
%icd_pcs_class(mix, Pr1 Pr2 Pr3 Pr4 Pr5 Pr6, distinct=0);
```

RESULTS

6 procedures were used for classification

17 records were read from MIX dataset:

5 of them were ICD-9-CM.

The mean number of codes per record was 3.40 (SD=1.82).

9 of them were ICD-10-PCS.

The mean number of procedure codes per record was 3.67 (SD=1.66).

3 of them were not classified

Some classified records had invalid procedure codes:

3 records had at least one invalid code.

Real ICD-9-CM Data, n=8,074,825

```
%icd_pcs_class(NIS,Pr1--Pr15, distinct=0);
```

RESULTS

15 procedures were used for classification
8074825 records were read from NIS dataset:
4995810 of them were ICD-9.
The mean number of codes per record was 2.49 (SD=1.96).
3079015 of them were not classified.
Some classified records had invalid procedure codes:
65038 records had at least one invalid procedure code.
1339 records had at least two invalid procedure codes.
100 records had at least three invalid procedure codes.
33 records had at least four invalid procedure codes.

It took 3 min to run the program.*

* Performance was evaluated on a laptop with Intel® Core™ i7-3520M 2.9GHz.

REQUIRED PARAMETERS

dsn = SAS data set with records coded as ICD-9-CM procedure codes or ICD-10-PCS, or their mix.

Ordervars= list of variables (procedures); it can be only one variable (Pr1), or a list (Pr1 Pr2 Pr7), or a range (Pr1--Pr10).

OPTIONAL PARAMETER

Distinct = specify if an additional analysis (number of valid distinct codes used, table of unique invalid codes) of procedure codes is required, there are two options: 0 and 1; default is 0, no additional analysis.

Conclusions

- SAS macro runs quickly on very large data sets with multiple dx per record.
- It can easily flag invalid codes (which are neither ICD-9-CM nor ICD-10-CM/PCS).
- It can easily flag records that incorrectly include both ICD-9-CM and ICD-10-CM/PCS codes for rejection or manual review.
- Diagnosis codes that are shared between ICD-9-CM and ICD-10-CM are allocated according to other codes on the same record and E code rules.

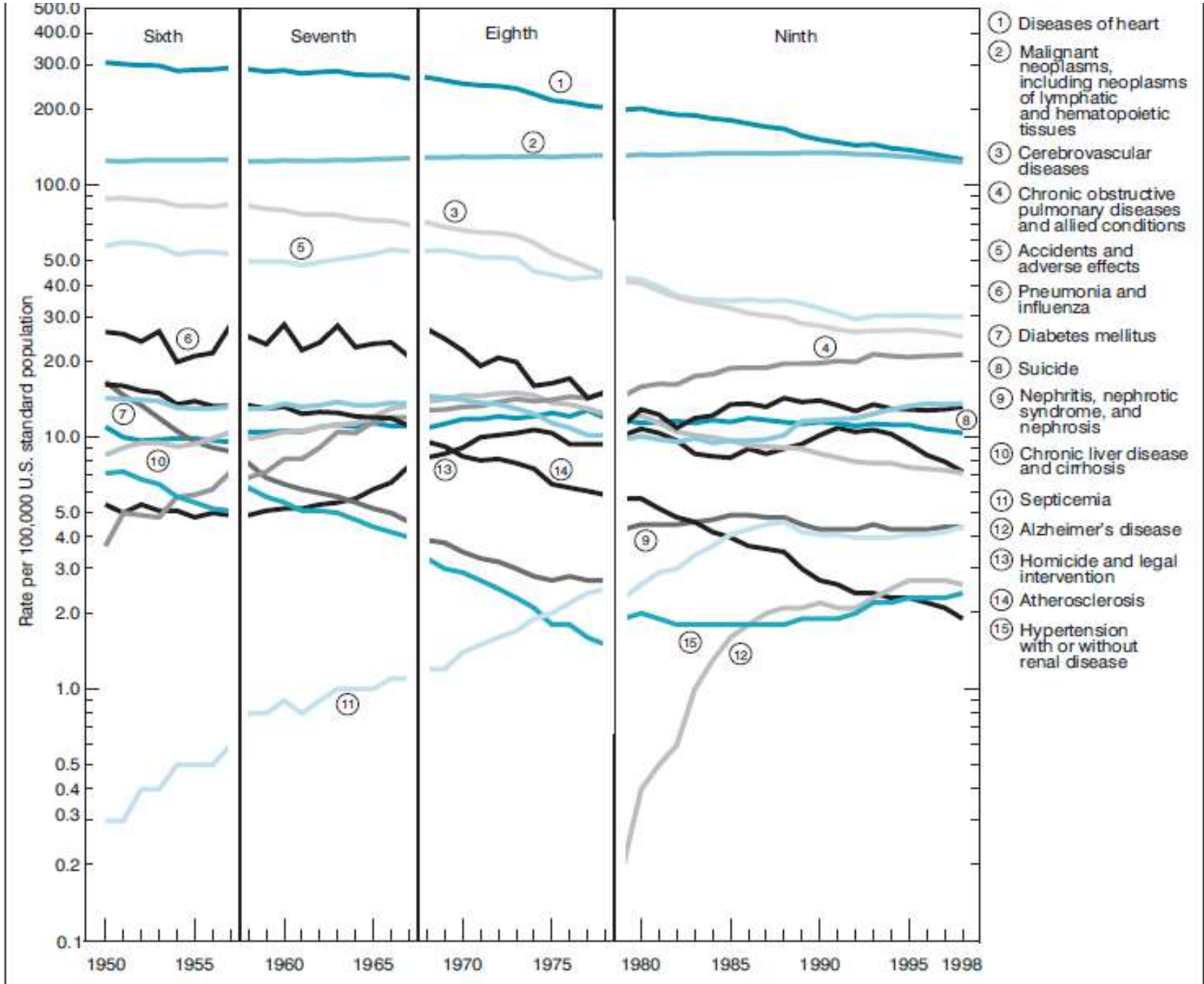


Figure 4. Age-adjusted death rates for the 15 leading causes of death: United States, 1950–98

Advantages of dual coded data

- Estimate comparability ratios for each cause of morbidity:
 - $C_i = D_i(\text{ICD-10-CM}) / D_i(\text{ICD-9-CM})$
 - where $D_i(\text{ICD-10-CM})$ is the number of records with cause of morbidity i classified by ICD-10-CM
 - and $D_i(\text{ICD-9-CM})$ is the number of records with cause of morbidity i classified by ICD-9-CM.
 - Ratio=1.00 indicates that the same number of records had morbidity i under both code sets, denoting no net effect of ICD-10-CM on that particular measure. This does not necessarily indicate that the cause was totally unaffected by ICD-10-CM, but merely that any increased allocations to morbidity i were completely offset by decreased allocations to morbidity i .
- Identify expected changes after 10/1/2015
- Identify potential errors in coding practice or code maps

Approaches to the problem

- CDC created dual coded mortality data set using 1996 US death certificates:
 - 1,838,968 death records were coded automatically
 - Only 13,703 records that could not be classified using automated systems were manually coded
- CMS created simulated dual coded data sets using all Medicare inpatient admission (Part A) claims from hospitals paid using Medicare Severity Diagnosis Related Groups (MS DRGs):
 - Used General Equivalence Mapping files with additional (proprietary 3M) logic to select the most plausible set of ICD-10-CM/PCS codes for each record
 - 10,009,934 admissions from 3,205 hospitals in FY 2014.
- Hospitals and health systems “natively” dual coded some of their own records for training and validation

Two dual coded data sets

Washington state

- 2,665 records with dual coded diagnoses and procedures
- 8 hospitals in 2013
- Exact method of dual coding varied by hospital
- AHRQ HCUP/QI program, WA State Dept. of Health

University of Wisconsin

- 2,191 dual coded records from July 2011, 2 hospitals
- 1,778 dual coded records from July 2012, 2 hospitals
- “assistive encoding software that suggested the appropriate codes”
- “outsourced or contract coders were used to backfill the ICD-9-CM coding gap”
- Fenton and Benigni

Case study: acute myocardial infarction for illustrative purposes only

- ICD-9-CM 410.x0, 410.x1 (some also include 410.x2, “subsequent episode of care” for an AMI that “received initial treatment” within 8 weeks, but those are not “incident” events)
- ICD-10-CM I21.xx, I22.xx (includes AMI “with a stated duration of 4 weeks [28 days] or less from onset”)
- **Washington CR=1.00 (n=39)**
- **Wisconsin CR=1.008 (127/126)**

Case study: asthma for illustrative purposes only

- ICD-9-CM 493.xx (asthma, excludes “chronic obstructive bronchitis,” 491.20-491.22)
- ICD-10-CM J45.xx (asthma, excludes “asthma with COPD, chronic asthmatic [obstructive] bronchitis, chronic obstructive asthma, J44.9)
- Problem: 493.2x (chronic obstructive asthma) maps to J44.x (other COPD)
- **Washington CR=1.00 (small numbers, with J44)**
- **Wisconsin CR=0.893 (excluding J44)**

Conclusions

- Although most of the available dual coded data sets are carefully protected under HIPAA, they may provide helpful insights for public health and quality measurement programs
- As for causes of death, comparability appears to be high for grouped conditions, but is often problematic at the individual code level (e.g., Fenton reports CR=0.41 for heart failure NOS, 428.1 versus I50.9)
- Contact Emily Sullivan, esullivan@nahdo.org, for more information
- Acknowledge AHRQ QI program and Washington State Department of Health (HCUP partner)