

# ACCURACY OF ICD-9 CODING FOR SICKLE CELL DISEASE (SCD) IN CHILDREN AND ADOLESCENTS: RESULTS FROM THE GEORGIA (GA) RUSH SURVEILLANCE PROJECT

Peter A Lane MD, Rodney Theodore MPH, Mei Zhou MS, Angela Snyder PhD

## RESEARCH OBJECTIVE

To study the accuracy of using ICD-9 codes within administrative claims data to identify individuals with Sickle Cell Disease (SCD).

## STUDY DESIGN

Retrospective study to evaluate the accuracy of the Registry and Surveillance for Hemoglobinopathies (RuSH) case definition in identifying true cases of SCD in administrative datasets. Probable cases required identification of  $\geq 2$  encounters with a SCD ICD-9 code (282.60-282.69; 282.41-282.42) plus  $\geq 1$  encounter with an ICD-9 or CPT code from a predetermined list of SCD-associated treatments, procedures, and complications (SCD-TPC).

| SCD-associated treatments   | SCD-associated complications                       |
|---|--|
| Hydroxyurea   | Chronic Renal failure/Proteinuria                  |
| Parenteral analgesics (Morphine, meperidine, hydromorphone, ketorolac, butorphanol) | Pneumonia, acute chest syndrome                    |
| Iron Chelators (deferasirox, deferoxamine)  | Pulmonary hypertension                             |
| Erythropoietin  | Stroke (ischemic or hemorrhagic), TIA, seizures    |
| Folic acid  | Intracranial bleeding                              |
|   | Priapism   |
|   | Iron overload                                      |
| SCD-associated procedures   | Gallstones/cholelithiasis, cholecystitis           |
| Red cell transfusion  | Avascular necrosis                                 |
| Red cell exchange   | Retinal disease                                    |
| Splenectomy   | Splenomegaly, splenic sequestration, hypersplenism |
| Cholecystectomy   | Leg ulcers   |
| Transcranial Doppler  | Dactylitis   |
|   | Osteomyelitis                                      |

Cases identified in administrative data were then compared to laboratory and clinical chart review data from Children's Healthcare of Atlanta (CHOA) to report on the accuracy of this case definition. Specific objectives were to evaluate accuracy in terms of 1) the number of SCD-coded encounters for each individual, 2) the addition of SCD-TPC to the case definition, and 3) the length of the surveillance period (1 versus 5 years of data). Cases with inadequate clinical data to confirm or exclude SCD were categorized as indeterminate.

## POPULATION STUDIED

A total of 1,998 children and adolescents with  $\geq 1$  encounter with a SCD ICD-9 code in three administrative datasets: Georgia Medicaid, CHIP, and State Health Benefit Plan claims data from 2004-2008 who were seen at least once at a CHOA facility and had medical records available for review.

## PRINCIPAL FINDINGS

SCD was confirmed in 1,763 (88.2%), excluded in 196 (9.8%), and indeterminate in 39 (2.0%) of the 1,998 children.

| Surveillance period (yr) | SCD ICD-9 codes | Patients | SCD confirmed | SCD excluded | Indeterminate | P-value* |
|--------------------------|-----------------|----------|---------------|--------------|---------------|----------|
| 5                        | $\geq 1$        | 1,998    | 1,763 (88.2%) | 196 (9.8%)   | 39 (2.0%)     | <0.01    |
| 1                        | $\geq 1$        | 1,454    | 1,386 (95.3%) | 57 (3.7%)    | 11 (0.8%)     |          |
| 5                        | $\geq 2$        | 1,851    | 1,735 (93.7%) | 96 (5.2%)    | 20 (1.1%)     | <0.01    |
| 1                        | $\geq 2$        | 1,373    | 1,333 (97.1%) | 33 (2.4%)    | 7 (0.5%)      |          |
| 5                        | $\geq 3$        | 1,746    | 1,693 (97.0%) | 46 (2.6%)    | 7 (0.4%)      | <0.01    |
| 1                        | $\geq 3$        | 1,248    | 1,231 (98.6%) | 14 (1.1%)    | 3 (0.2%)      |          |

\*P-values calculated by chi-square tests. P-value tests whether the proportion of those confirmed to have SCD using 5 years of data differs significantly from the proportion of those confirmed to have SCD using 1 year of data (i.e. compares 88.2% to 95.3%).

The likelihood of having SCD significantly increased with an increasing number of SCD ICD-9 codes during both time periods. The 1-year period had better accuracy, but more missed cases. For the 5-year surveillance period, the accuracy of  $\geq 2$  SCD ICD-coded encounters (93.7%) increased to (97.0%) with the addition of  $\geq 1$  encounter with SCD-TPC, but the number of missed cases increased from 28 (1.6%) to 251 (14.2%).

Table 3. Effect of number of SCD ICD-9 codes and inclusion of ICD-9 and CPT codes for SCD-associated treatments, procedures, and complication to RuSH surveillance case definition (2004-2008)

| Case Definition  | Patients (n) | SCD confirmed | SCD cases missed |
|--|--------------|---------------|------------------|
| $\geq 1$ SCD ICD-9 code  | 1,998        | 1,763 (88.2%) | 0                |
| $\geq 2$ SCD ICD-9 codes                                       | 1,851        | 1,735 (93.7%) | 28 (1.6%)        |
| $\geq 2$ SCD ICD-9 codes & $\geq 1$ ICD-9/CPT code for SCD TPC | 1,558        | 1,512 (97.0%) | 251 (14.2%)      |
| $\geq 3$ SCD ICD-9 codes                                       | 1,746        | 1,693 (97.0%) | 70 (4.0%)        |

## CONCLUSION

Use of administrative claims data to identify children and adolescents with SCD based on ICD-9 coding can be useful to health services researchers, but it has some limitations depending on the research question. Accuracy of identification correlated directly with the number of SCD-coded encounters and indirectly with the length of the surveillance period. The addition of  $\geq 1$  encounter with SCD-TPC to a case definition of  $\geq 2$  encounters with SCD ICD-9 codes minimally improved specificity but resulted in a large number of missed cases, more than the simpler definition of  $\geq 3$  SCD-coded encounters.

## IMPLICATIONS FOR POLICY OR PRACTICE

Health services and outcomes research in SCD has been limited by inadequate population-based surveillance and reliance on ICD-9 codes to identify individuals with SCD. This study supports an administrative case definition that includes 3 or more separate ICD-9 codes in order to identify SCD with a high degree of accuracy and suggests it can be useful to assess quality of care. However, using administrative data, by design, excludes individuals who have less health care utilization and are therefore less costly. Those limitations must be weighed against the benefits of using the data in future health services research.