

Conversion of Nationwide Health Insurance Claims Data into the Common Data Model in Korea Data Validation with 7.6 Million Cancer Patients' Healthcare Utilization Data for the Past 11 years

Yeunsook Rho

INTRO

- Due to i) differences in data language and structure among major healthcare data holding organizations; and ii) strong privacy protection act and other legal constraint, it has been difficult to actively use healthcare data in Korea.
- This study statistically verifies whether the converted CDM data produces consistent outcome with existing analysis of original data.

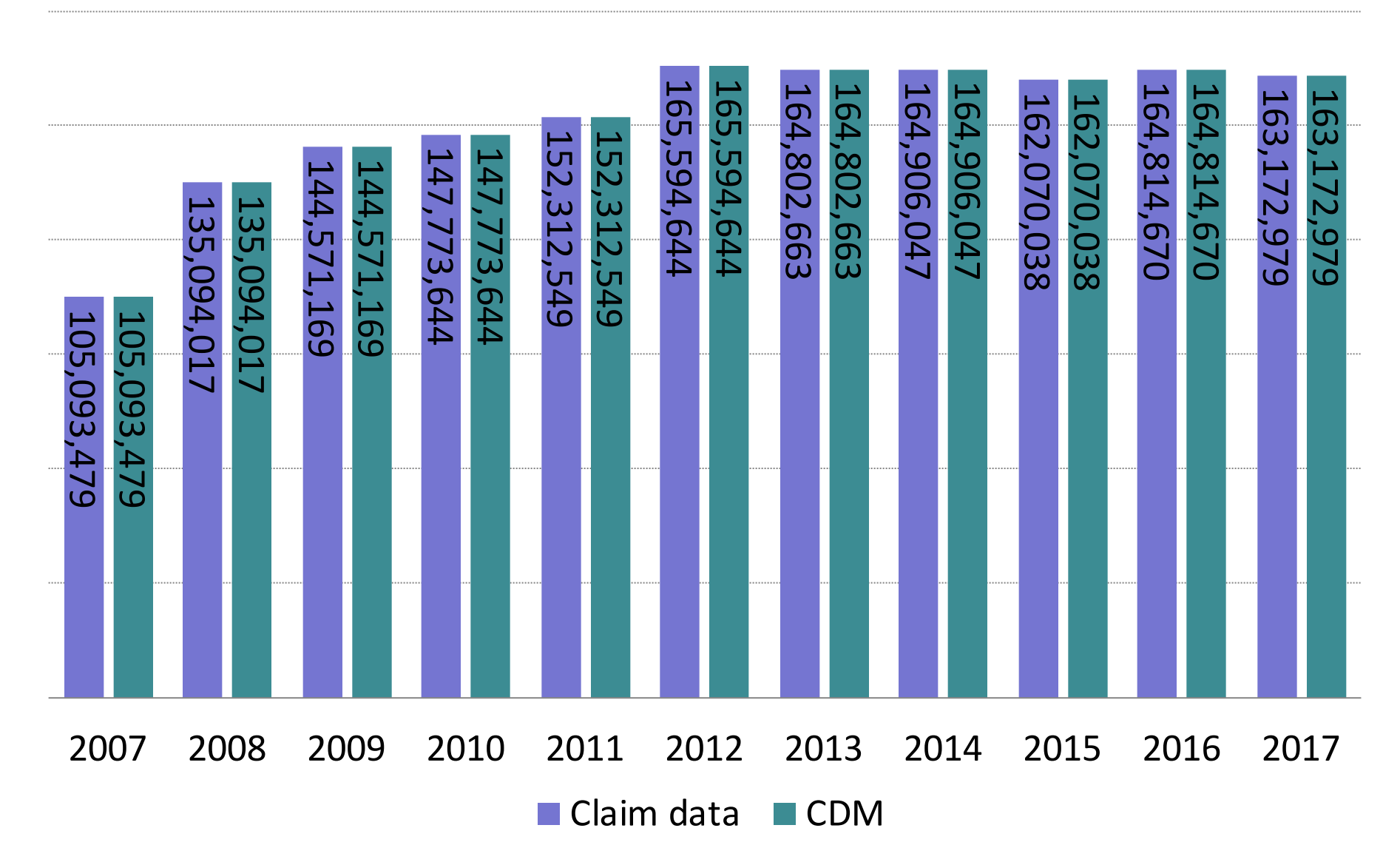
METHODS

- Some 7.6 million cancer patients' healthcare utilization data of 11 years were extracted from HIRA^{1*} claims data (2.6 TB) and the data were converted to OMOP CDM.
- *Health Insurance Review & Assessment Service
- Statistics were produced for different perspectives (diagnosis, procedure, and drugs) of the two datasets.
- Alongside, crude rates of major cancers (colon, stomach, lung, liver, breast and thyroid cancer) based on the CDM data were calculated to compare against National Cancer Registry Statistics

RESULTS

- The original claims data and CDM data showed identical values in the number of claims and patients by year.
- The number of claims and the order of most frequently claimed items in diagnosis, procedure, and drug also matched between the two datasets.
- Also, cancer crude rates of HIRA CDM data showed similar tendency with National Cancer Registry statistics.

Figure 1. Comparison of Number b/w HIRA Claims data and its CDM Data: Total Claims (Unit: Claim)



In Korea, The converted CDM data and original nationwide claims data had the identical statistic values in number of claims and patients at different dimensions.



Figure 2. Comparison of Number b/w HIRA Claims data and its CDM Data: Diagnosis_Colon Cancer (Unit: Claim)

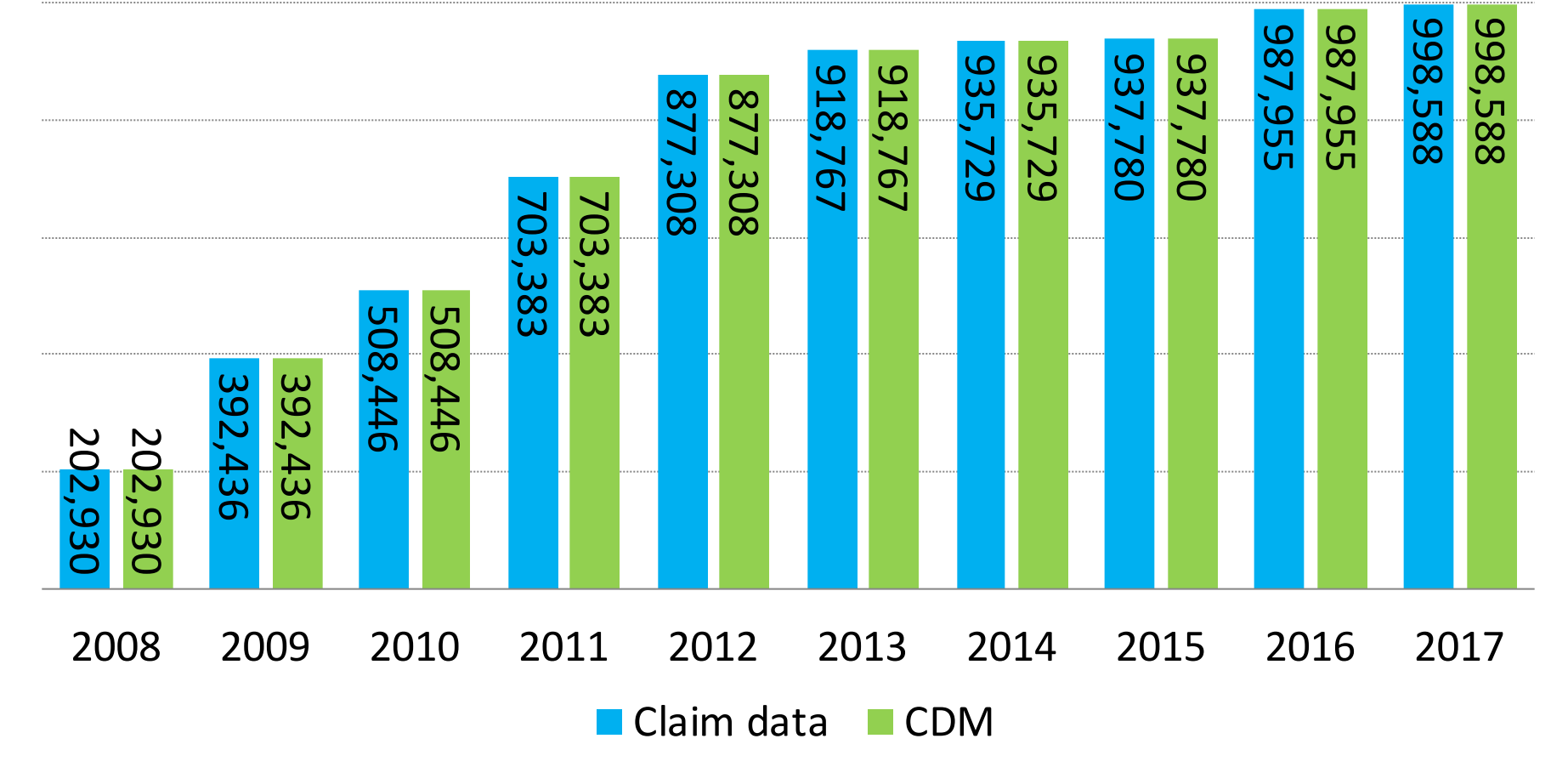


Figure 3. Comparison of Number b/w HIRA Claims data and its CDM Data: Procedure_Colonoscopy (Unit: Claim)

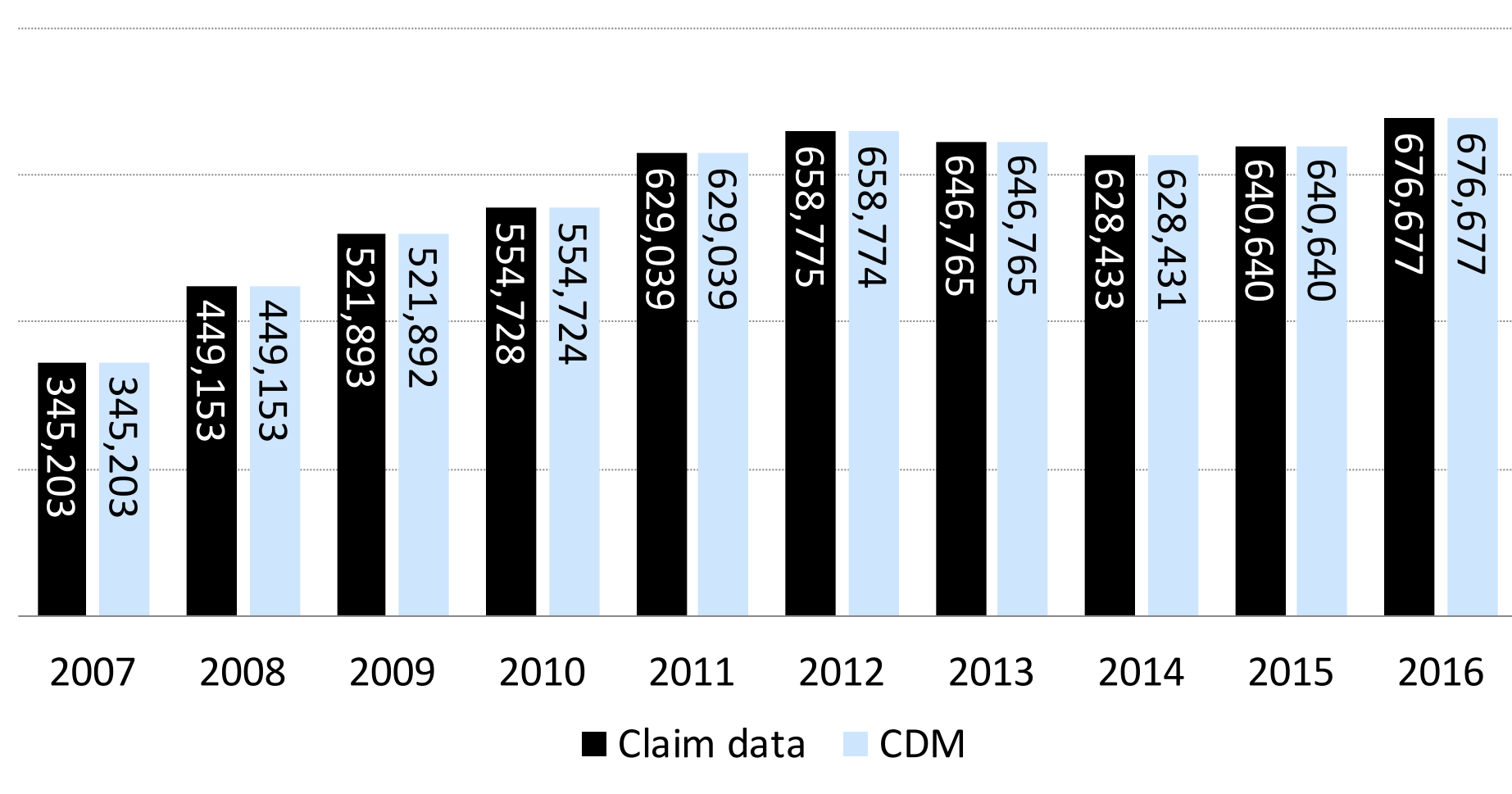


Figure 4. Comparison of Number b/w HIRA Claims data and its CDM Data: Drug_Iressa (Gefitinib 250mg) (Unit: Claim)

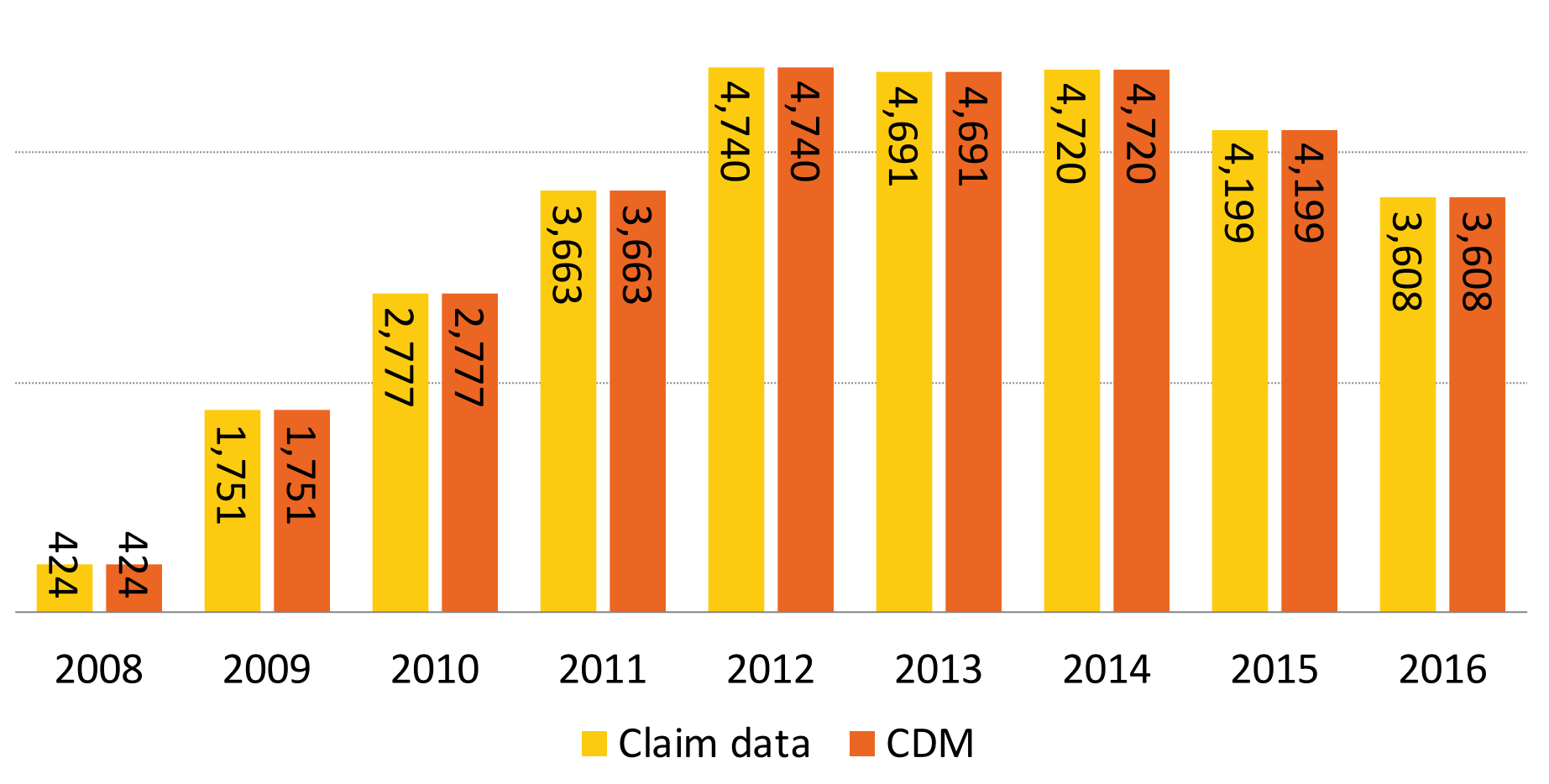
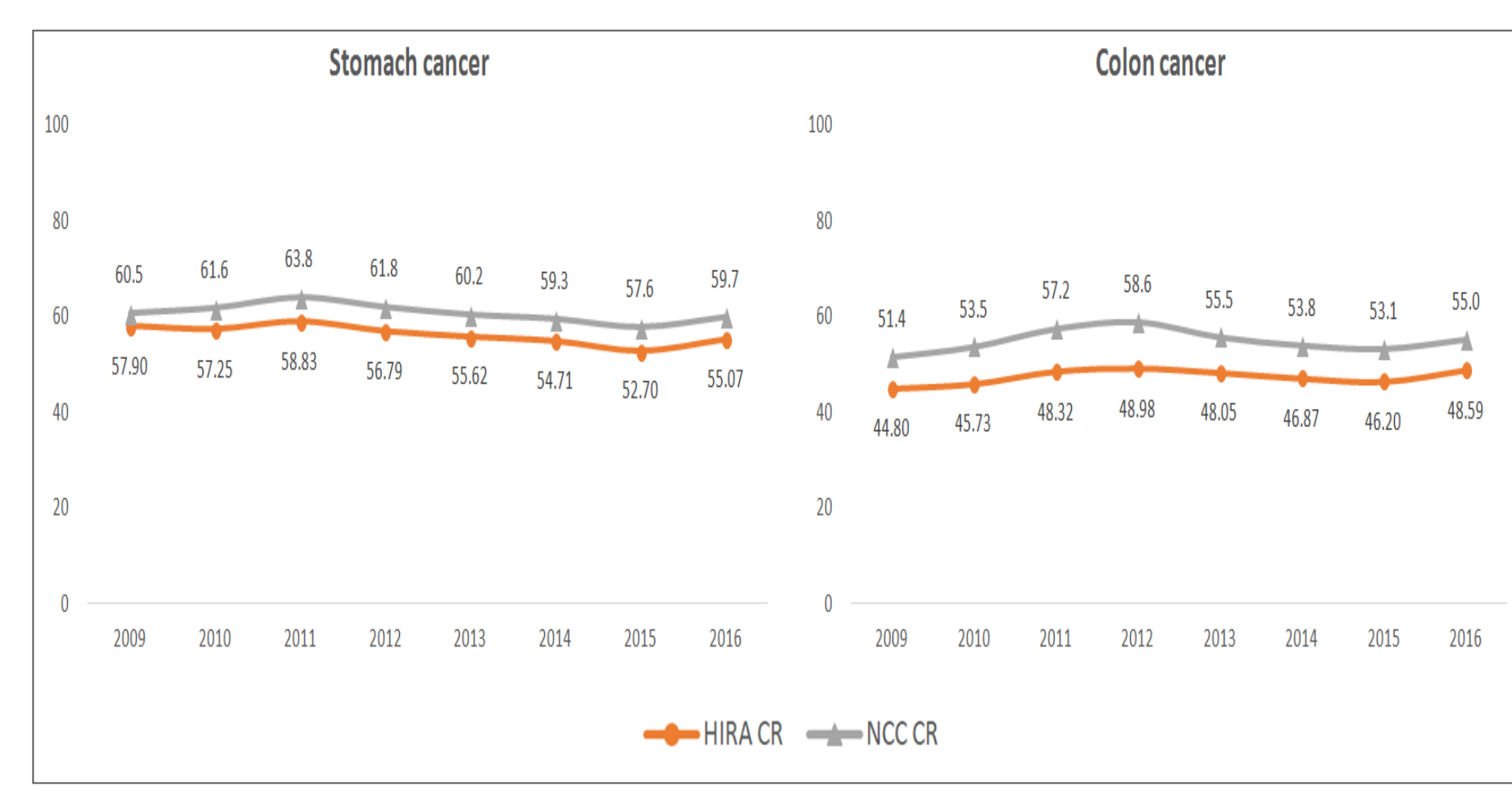


Figure 5. Trend Changes in the Crude Rates (CRs) of Stomach and Colon Cancers (Unit: Rate per 100,000)



Yeunsook Rho, Jiwoo Kim, Jimyung Park, Jaehyeoung Cho, Youjin Park, Seng Chan You, Yejin Son, Raewoong Park, Yun Jung Heo

