

AI vs Human Coding in DRG Severity: A Real-World OMOP CDM Study (Belgium)

AI-assisted coding identifies clinically valid missed diagnoses, increasing DRG severity and hospital capacity by +15% — but requires human validation to avoid errors.

PROBLEM & METHODS

Why this matters

DRG & SOI drive hospital financing in Belgium

Manual coding is:

- Time-consuming
- Variable
- Incomplete

AI can detect diagnoses from clinical text

But impact on DRG/SOI is unclear

Study Design

18,515 hospitalizations (2025, CHU Liège)

Compared:

- Human-coded ICD-10
- AI-coded (Solventum 360)

Derived:

- APR-DRG
- Severity of Illness (SOI)

Validation

Selected 2,563 cases where:

- AI → higher SOI

Reviewed by:

- 25 professional coders

Checked:

- Real diagnosis vs hallucination

OMOP Innovation

Stored different DRG/SOI in:

- OMOP OBSERVATION table

Linked to hospital stay

Enables:

- Multi-source comparison
- Full provenance tracking

- Future:
- AI benchmarking
 - Multi-center OMOP studies
 - ICD-11/ SNOMED integration

RESULTS & IMPACT

Key Findings

→ identical DRG (AI vs human)

10,312 / 18,515

→ In higher AI-SOI cases:

→ many AI diagnoses were **clinically valid**

Impact

+15% increase in effective justified capacity

Equivalent to: +7.67 hospital beds

Mean LOS: 5.79 days

Driven by:

- Additional **secondary diagnoses**
- Improved SOI classification

Interpretation

AI alone → not reliable

Human alone → incomplete

Best approach: Hybrid Human + AI

Takeaways

AI improves **coding completeness**

Human ensures **clinical validity**

OMOP enables:

- Transparent comparison
- Scalable evaluation

