TITLE: Accuracy of DRG Relative Weights Calibration using Machine Learning versus Standard HSRV Methodology

Introduction

The standard methodology for calculating Diagnosis Related Groups (DRG) relative weights is based on the average costs associated with treating patients assigned to each DRG. This methodology involves estimating costs using a national 'cost-to-charge' (CCR) ratio for defined cost centers and standardizing these estimations to account for systematic differences across hospitals, such as geographic variations in wage levels, type of hospital (small, medium, large, teaching, etc.) and costs of living. The relative weight for each DRG is then determined by dividing the average standardized cost for discharges assigned to that DRG by the average standardized cost for all discharges, reflecting the relative resource intensity required for different types of care.

Methods

The "Hospital-Specific Relative Value (HSRV)" methodology tailors the calculation of relative weights, accounting for differences in case-mix between individual hospitals, in an iterative regression. Adjusting the hospital CMIs with newly calculated weights in each iteration is performed until convergence is achieved i.e. variance of CMIs between iterations is minimal. All methodologies employ statistical methods to eliminate LOS and/or financial outliers and adjust for cases that may skew results. With severity of illness systems (APR-DRG, IR-DRG, AR-DRG) monotonicity adjustments of the obtained relative weights are further applied to obtain increasing values with more severe groups of patients within each DRG split by severity of illness.

In a machine-learning (ML) approach using generated data, relativities between DRGs were estimated using raw charges instead of estimated costs, without outlier exclusion. The adjusted R^2, variance and sensitivity (AUC) of the relative weights obtained using machine learning were compared with relative weights derived using the HSRV standard methodology.

Results

The study aims at discarding tedious cost/charge collection and validation steps from the legacy methods without significant accuracy loss, when predicting relative weights for inpatient encounter reimbursement.

Discussion/conclusions

Extensive cost studies are needed to obtain the input data for determining the relative weights using the common methods described above. The calculation may be biased by several factors including the size of the sample, lack of randomization when including available hospital data from an unrepresentative cohort of hospitals, inaccurate assumptions for cost estimates or allocative cost algorithms rather than activity-based alternatives. Using a machine learning approach may reduce the recalibration project duration and may offer qualitatively similar relative weights when compared to the HSRV methodology.