2008 National Health Insurer Report Card: Statement of methodology

The purpose of the American Medical Association's National Health Insurer Report Card (NHIRC) is to provide physicians and the general public with a reliable and defensible source of critical metrics concerning the timeliness, transparency and accuracy of claims processing by the health insurance companies that are responsible for paying these claims. This report card includes results for Medicare and seven commercial health insurers' (payers) related to their payment timeliness, accuracy, transparency of contracted fees and payment policies, compliance with generally accepted pricing rules, and denials. These metrics will be used to identify potential payer and physician practice deficiencies with a goal to improve the overall efficiency of billing and collections, thereby reducing the cost to both the physician practice and payer. To accomplish this task, the AMA used its best efforts to ensure that the methodologies employed to obtain and analyze the data are sound and defensible and the process is transparent such that all interested parties are able to understand and duplicate the study.

Health insurer (i.e., payer) selection

The payer was identified from the payer name field on the electronic remittance advice (ERA[1]). Selection of payers was based on their national reach as evidenced by their large numbers of national commercial enrollees. Medicare and the following seven commercial payers were selected for this initial report card.

- Aetna
- Anthem Blue Cross and Blue Shield
- CIGNA, Corp.
- Coventry Health Care
- Health Net, Inc.
- Humana, Inc.
- UnitedHealthcare

[1] See the next section (Data Source) for a description of ERA.

Data source

The AMA's NHIRC will report 14 metrics, which can be categorized into five groups.[2] The data used to produce those metrics come from two sources. The major source was used to obtain information on four of the five groups: Payment Timeliness, Accuracy, Compliance with Generally Accepted Pricing Rules and Denials. This data source is a data base maintained by National Healthcare Exchange Services (NHXS), a company located in Sacramento, CA. NHXS is an application service provider that offers a pricing and payment audit system. The NHXS' master data base from which the NHIRC results were drawn includes over 5 million services billed on over 3 million claims between the second half of 2007 and the first quarter of 2008. Claims come from 20 states, over 7,500 practicing physicians representing 18 specialties, and 195 practices. NHXS reports data from claims transactions based on the Federally mandated Health Insurance Portability and Accountability Act of 1996 (HIPAA) electronic standard transactions. The technical references for these transactions are the electronic remittance advice, (ERA) (HIPAA ASC X12 835 Health Care Claim Payment/Advice Transaction), which is submitted to a physician in response to the receipt of an electronic claim submission (HIPAA ASC X12 837 Health Care Claim—professional transactions.)

Information on the fifth group of metrics (i.e., transparency of contracted fees) was obtained directly from the payers' Web sites.

Payment timeliness

Sample size and stratification

The definitions below will add clarity when referencing each sample set identified in the Payment Timeliness study:

- Master data base: over 3 million claims, including over 5 million records (i.e., lines on claims) submitted between July 1, 2007 and March 31, 2008.
- Sample universe: 600,000 records extracted from the Master Data Base representing only the eight payers evaluated for this study.
- Final sample: approximately 10,000 records randomly selected from the sample universe for each of the payers (except Health Net[3]).
- Test sample: 10 random samples of 350 records from each payer obtained from the final sample.

Due to the large size of the sample universe, a subset of records was obtained to allow for a more critical review of outliers and to improve our ability to conduct significance tests both within and between payers. This subset consisted of 10,000 records randomly selected for each of the payers (except Health Net) and was used as our final sample. To build confidence in the final samples and to gain a better understanding of the distribution, a validation analysis was conducted to ensure that each of these final samples represented the sample universe. For this exercise, a subset was obtained from each payer's final sample; we refer to the smaller samples as the test samples. To determine initial statistical power and sample size for this step, a study of variability (i.e., standard deviation for the sample universe) was performed. The first sample size analysis indicated a standard deviation of approximately 6.5, which suggested the appropriate initial sample size should be 350 records. The next step was to extract 10 random test samples of 350 records from each payer and perform an ANOVA[4] across these samples to ensure they represent data from the same population with respect to the metrics to be analyzed in the study. This gave us confidence that the 10,000 record samples were representative of the sample universe.

---

[2] These groups (and the corresponding metric numbers on the Report Card) are: a) Payment Timeliness (1-3), b); Accuracy (4-5); c) Transparency of Contracted Fees (6-9); d) Compliance with Generally Accepted Pricing Rules (10-11); and e) Denials (12-14).

[3] We were able to obtain only 5,500 records for Health Net due to the market limitations of NHXS at the time of the request.

[4] ANOVA (analysis of variance) is a statistical technique used to test for differences between two or more groups; in this case, the groups in question are the 10 random test samples of 350 records.

Once the final samples were validated, i.e., the first remittance response time metric median values yielded by the test samples were not statistically different from those yielded by the sample universe, we used them to conduct the statistical analysis to produce the final metrics used for this study.

The data for Medicare was used as a benchmark for central tendency and variability measurements and were compared to the other payers' statistics. Once the Medicare data was validated, i.e., the median and IQR values were not statistically different from those yielded by the final sample, we felt comfortable that the data for the remaining payers' sample sets were also validly selected, since the methodology for processing and selection of the records were the same for each payer.

Accuracy

For an electronic remittance advice (ERA) to be considered accurate, the allowable amount the payer reported to calculate payment must match the physician's contracted payment rate with the payer. We define "contracted payment rate adherence average" as the number of records yielding a zero difference between the reported allowed amount and the physician's contracted payment rate as a percent of all records.[5] Conversely, an accurate ERA is defined as one where the payer's reported allowed amount is equal to the physician's contracted rate, leaving their difference equal to zero. Contract allowable amounts were entered into the sample universe based on NHXS obtaining either the contracts or the contracted fee schedule from the physician when these were available; otherwise they were obtained from the payer. For Medicare, fee schedule transparency was not an issue as the Medicare Fee Schedule allowable amounts are published each quarter.

The contracted payment rate adherence average metric accounts for the potential inconsistency between a single payer's allowed amount (contracted fee schedule) for the same physician, AMA Current Procedural Terminology (CPT®)[6] code, product type and fee schedule effective period. Data were pulled for a two month period from February and March 2008 since, traditionally, contract fee schedules do not change during these months.[7] NHXS pulled 100% of all claims reported for the eight payers, which resulted in an initial file containing 1,607,728 records. However, only certain records were selected in an attempt to control for modifiers, place of service, CPT code and physician contracted amount. Specifically:

- Only modifiers 26 (professional component) and TC (technical component) were included in the analysis because these two modifiers are typically associated with specific allowed amounts. All other modifiers and all modifiers in positions beyond position 1 were excluded.
- Place of Service indicators were restricted to those identifying the following:
    - Medical Practice (11)
    - Inpatient Hospital (21)
    - Outpatient Hospital (22)
    - Emergency Room (23)
    - Skilled Nursing Home (31)
    - Nursing Facility (32)

---

[5] The magnitude of the difference in the allowed amount is not considered in this metric because any difference in the allowed amount from the physician's contracted payment rate was considered to be a deviation from a proper ERA.

[6] CPT is a registered trademark of the American Medical Association.

[7] In the majority of cases, contracts undergo review for fee schedule modification every calendar quarter. Data for the first month of each quarter (e.g., January) may not be appropriately updated by January 1 and include fees from the last month of the prior quarter. Therefore comparisons may be inappropriate due to the additional time necessary to update any fee schedule modifications. By using the second and third month of any quarter (e.g., February and March), there is enough time to update fee schedule modifications in the previous month, which generally avoids inconsistencies in fees.

- We excluded the following CPT codes: anesthesia (0000-9999) and the new CPT tracking codes (ending with a T or an F).
- Finally, only contract amounts > $1 were included.

These sample selection restrictions left us with a total of 1,186,905 records representing the eight payers, ranging from 720 for Health Net to nearly 900,000 for Medicare.[8] In order to control for the possibility that major code categories might influence the results, this sample was divided by the following procedure codes:

- Surgical procedures – 10000 through 69999
- Radiological procedures – 70000 through 79999
- Pathology and Lab procedures – 80000 through 89999
- Medicine procedures – 90000 through 99200
- E/M procedures – 99201 through 99499

For each payer, the analytical process involved subtracting the expected contract allowable amount from the payer-reported allowable amount and arranging these values by procedure code category. The number of records with a zero difference was then divided by the total number of records for that category. The result is a percent that represents our contracted payment adherence average metric.

Transparency of contracted fees and payment policies on payer Web sites

The methodology for producing these measures consisted of simply determining whether the contracted fee schedule, payer proprietary code edits and medical payment policies were accessible on each payer's Web site.

Compliance with generally accepted pricing rules

Sample size and stratification

We selected a sample of records from NHXS' master database that were submitted to health insurers by physicians representative of most specialties[9] during the 3rd and 4th quarter of 2007. This selection criteria resulted in 3,763,814 records, ranging from 11,124 for Coventry to 3,026,809 for Medicare.[10]

A claim edit is defined as a payer applied adjustment that reduced the payment (allowed amount) of the claim line to $0. Edits consist of coding and payment rules that result in a denial of payment for a reported service. An example of an edit is a "denial of a service as inclusive of another reported service."

To determine the impact of claim edits on the Medicare and commercial payers' claims processes, the following types of edits and the sources they were based on were entered into the NHXS claim editing engine:

- Edits based on AMA CPT codes, guidelines and conventions; National Correct Coding Initiative (NCCI); Centers for Medicare & Medicaid Services (CMS) Publication 100-04; and the American Society of Anesthesia (ASA) Relative Value Guide.
- Payer-specific proprietary edits that reflect the specific clinical edit policy, along with additional edits other than those based on CPT, NCCI and Medicare payment rules applied at the prerogative of the payer.

---

[8] The number of records for each payer is as follows: Aetna, N=78,650; Anthem BCBS, N=29,497, CIGNA, N=28,072; Coventry Health Care, N=4,919; Humana, N=11,833; Medicare, N=898,672; UnitedHealthcare; N=134,542.

[9] Anesthesiology, Cardiovascular, Dermatology, Emergency Medicine, Family Practice, Internal Medicine, Multispecialty, Neurology, Neurosurgery, Ophthalmology, Orthopedics, Pathology, Pediatrics, Pulmonology, Radiology, Surgery and Vascular Surgery.

[10] The number of records for each payer is as follows: Aetna, N=186,570; Anthem BCBS, N=75,031; CIGNA, N=72,320; Coventry Health Care, N=11,124 ; Humana, N=40,020; Medicare, N=3,026,809; UnitedHealthcare; N=351,412.

Of the 3,763,814 records analyzed, claim edit adjustments were made to 90,221 records. For each payer, we present the percentage of records (1) reduced by edits and (2) where the source of claim edits applied by the payer were based on one or more of the following: AMA CPT, NCCI, CMS Publication 100-04, ASA Relative Value Guide, or payer proprietary edits.

Denial analysis

The purpose of the denial analysis is twofold. First, denials were calculated as a percentage of records in the sample. Second, we conducted a rank-order analysis of the denial reason and remark codes given by the payers on the HIPAA 835 X12 standard transaction for the payment adjustment.

Denials were the second major analytical component obtained from the first payer ERA that was sent to the physician. For the purpose of this study, we defined a denial as any case where the payer allows the physician-billed charge but the actual payment is zero. In each of these cases, the ERA was accompanied by reason and remark codes identifying why the payer denied payment. Specifically, denial codes are reported in the ERA using three levels of indicators. Level one consists of five Claim Adjustment Group Codes (CAGC), level two consists of 258 reason codes, of which 190 were available for use as active during the data period, and the third level consists of approximately 675 remark codes, which may be associated with specific reason codes. The five adjustment group codes that make up the first level are 1) CO: Contractual obligation, 2) CR: Correction and reversals, 3) OA: Other adjustment, 4) PI: Payer initiated reductions, and 5) PR: Patient responsibility. Note that more than one CAGC, reason code and/or remark code can appear within a single claim line.

Because of the large number of possible combinations of the three levels (i.e., CAGC, reason, remark), 9,385,647 records were reviewed with date of services from 3/1/2007 to 3/10/2008. For the eight payers, this represented a total of 574,591[11] records associated with denials containing reasons codes ranging from 193 for Health Net to 475,566 for Medicare, and 298,651 records associated with remark codes, ranging from 59 for CIGNA to 252,882 for Medicare. Note, of course, not every reason code has an associated remark code.

To examine whether there was a pattern of use between these three levels (i.e., CAGC, reason and remark codes) a Pareto analysis was conducted to determine which denial code group combinations are associated with the majority of denials reported in the first ERA. This was done for each payer and then by both payer types (i.e., all private payers and Medicare). We present Pareto graphs[12] that show, in descending order, the number of data points associated with a particular value as both raw volume and a percent of total denials. These graphs represent the relationship between reason code and group type, reason code and payer and remark code and reason code. Finally, we examined whether there was a pattern of use of the denial codes within each payer. This analysis sheds light on the logic pertaining to the use of these three denial code groups by examining the descriptions for each group and determining the appropriate use of the combinations of code types. Ratios of total denials are illustrated using these Pareto graphs.

[11] Number of records for each payer are as follows: Aetna, N=43317; Anthem BCBS, N=11546, CIGNA, N=9060; Coventry HC, N=590; Health Net Inc., N=193; Humana, N=4142; Medicare, N=475566; UnitedHealthcare; N=30177.

[12] A Pareto graph is a representation of the proportion ranking of a given set of data.