Data quality in coding of diagnoses

Per Hansson, Head of Department, Patient Classification, Stockholm County.
per.hansson@sll.se
Lennart Green, senior consultant, System Futura AB.
green.systema@telia.com

Abstract

In the Stockholm County, responsible for the health care delivery system for almost 1/5 of the Swedish population, the coding of 400-650 medical records have been reviewed by coding experts annually over the last 5 years. The sample of records to be reviewed has been focused on problem areas where miscoding might be more expected to occur. The finding of these reviews is that almost 15-25 % of the primary diagnoses are coded incorrectly. More than 10% of the cases were assigned a wrong DRG. During the last few years more than 100 medical secretaries in the Stockholm County have passed a 20 days training course in coding of diagnoses. Nevertheless the coding still in most medical departments rests with the doctors who are unwilling to transfer the coding to the trained secretaries. In order to compare the quality of the coding between medical doctors and specially trained secretaries a study was conducted based on 210 medical records from six clinical departments (two from each surgery, orthopaedics and internal medicine).The system presented at the PCS-meeting at Budapest by dr Michael Wilke was used to measure the quality of the coding. The outcome of the study indicated that the secretaries had a score which was 20% better than the score of the medical doctors.

Background

Reimbursement to hospitals for inpatient care in the Stockholm County, responsible for all health care services in the Stockholm area (approximately 1.8 million inhabitants), has since the early 1990-ties been based on prospective DRG-costs. During the first two years productivity based on DRG-points divided with total costs increased with almost 10 % per year. Many questions were asked about the reason for this dynamic change. Was it due to the introduction of DRGs and/or the cancellation of the global budget system with fixed budgets for the hospital departments? Researchers in 1997 approached the problem and it was evident that both factors played a major role.
But it was also obvious that the average weight per case in DRG-points increased as the clinicians got a better understanding of the DRG-system. This envisaged DRG-creep had both legal and illegal grounds. The legal creep was due to a better understanding of the DRG-system and thereby a better and more correct coding. The illegal creep which was estimated to be between 30 and 40 % of the total creep was caused by incorrect not always deliberate coding of diagnoses.

Why coding reviews?

In order to come to grip with the problem of the DRG-creep the County Council of Stockholm decided to perform yearly audits of the coding. Other reasons for doing reviews of the coding were

- To find out if additional training is needed?
- To improve the coding criteria (coding manual)
- Correct and fair reimbursement

To describe how these audits were performed and measures taken based on the results of the audits is the topic of this paper.

Prerequisites

All hospitals, private or public, are obliged since more than 10 years, to deliver basic data (i.e. including diagnoses, procedures and administrative data) about all patients treated to an encrypted central database. Knowing the specific citizen number all healthcare data about a specific patient is possible to retrieve after permission of the patient. An encrypted dummy system makes it possible to know about all healthcare delivered to an individual patient without knowing the identity of the patient.

It is a tradition in the Scandinavian countries that coding of diagnoses is performed by the medical doctors with a very limited education in the ICD-coding system. Many Swedish medical secretaries with a 1 or 2 years post high school education in medical informatics are very rarely, even with a more extensive training, performing the coding. Most doctors are not often aware of what skills and knowledge are needed to perform a correct coding of diagnoses. A question at issue that has been discussed over the last few years is how the quality of coding would be affected if the medical secretaries were allowed to take more responsibility for the coding.

Coding Reviews procedures

Coding reviews have been conducted during the last five years (2000-2004) in the Stockholm County. Between 400 and 650 case records have yearly been reviewed. The reviews of the medical case records have been carried out by doctors from outside Stockholm with a deep knowledge in the coding criteria recommended by WHO and the additional Swedish criteria. Besides the quality improvement project that will be discussed later in this paper all other reviews were carried through in the same pattern. The review process thus involves five basic steps

- Problem definition /Sampling
- Retrieval of sampled case records
- Audit by two or three auditors
- Feedback
- Documentation

Sampling
The method for selecting case records to be reviewed could either be random sampling or using a problem oriented approach. The method used in Stockholm is a combination of the two methods. First the comprehensive and extensive database of all inpatient care delivered in Stockholm was used to find areas where hospitals or departments with similar patient groups showed a different DRG panorama. For these departments and DRGs a random sample of case records were picked out for the audit.
Retrieval of case records
The clinical departments were presented a list of case records to be audited. They were instructed to copy specified documents with relevant information for the review of the coding.
Audit
The case records first were reviewed by two separate auditors. When the auditors came out with a difference in their coding they sat down to discuss the coding. If they were not able to reach a common understanding the coding problem was referred to another auditor the umpire.
Feedback
The case records with tentative coding protocols for each case were sent back to the clinical departments together with a request for comments. After reviewing the comments the coding protocols were finalized.
Documentation
A final report with the results was compiled and distributed throughout the county via the web.

A seminar was arranged in order to discuss the findings. The criteria in the coding manual in some cases were amended.

Methods for describing coding errors
The auditors used a coding system, described below, to define if the codes were correct or what type of error they had found.

Manual for error coding		
Principle diagnoses	-correct diagnosis, correct code	
	-correct diagnosis, wrong code	
	-correct primary diagnosis coded as secondary diagnosis	
	-wrong diagnosis	
Secondary diagnosis	-correct diagnosis, correct code	
	-correct diagnosis, wrong code	(occurs/diagnosis)
	-wrong diagnosis	(occurs/diagnosis)
	-diagnosis missing	(occurs/diagnosis)

This coding of course could be discussed as not being too fair and not differentiating errors in terms of gravity. The quality improvement project, which will be presented later in this paper, used this system together with dr Wilkes scoring system which perhaps differentiates errors in secondary diagnoses in a better way.

Findings
The findings accounted for in this paper relates to the three first reviews focused on the inpatient services whereas the following reviews were more special in their focus. One was focused on one day surgery and one on obstetrics and childbirth only.
In this table the percentage of coding errors is presented together with the related change in assigned DRG.

Coding errors

Year	% coding errors errors			
	Primary diagnosis		Secondary diagnoses	DRG-changes
	Major	Minor		
1	15%	2%	45%	26%
2	21%	3%	67%	21%
3	12%	7%	63%	14%

Minor error in the primary diagnosis indicates an error in the 4th or 5th position of the ICD-code.
The number of errors in primary diagnoses has not changed very much but the errors in secondary diagnoses increased over the years which are supposed to be an effect of a deeper understanding of the DRG-system where some secondary codes might affect the DRG. What is interesting to notice is that changes in assigned DRG decreased significantly over the years.

As a result of these findings a two day training course for doctors and medical secretaries in coding was developed and offered to the clinical departments. The hospitals also started their own in-house training programs. The coding manual was also updated in terms of the usage of some secondary diagnoses specially those that might upgrade the DRG.

One can also notice that the number of diagnoses coded per case increased. The numbers shown in the table below relate to the total database with more than 200 000 cases per year. It is also of interest to notice that university clinicians have gained a better understanding of correct coding.

Even if the number of diagnoses per case has increased with more than 0.3 diagnoses over a 4 year period there is still missing another 0.5 diagnoses per case according to the reviews.

Type of hospital	Year			
	1	2	3	4
University hospitals	2.03	2.15	2.33	2.51
Other emergancy hospitals	2.19	2.25	2.26	2.32
Total	2.09	2.18	2.30	2.41

Quality improvement project

To increase the correctness in the registration of diagnoses is of major interest because it will give a much better knowledge of what is happening in the health care and it will also improve the use of reimbursement systems based on DRG.
Even if the clinical doctors' interest in correct coding has increased over the years there still remains an insufficient knowledge and training with the doctors in coding of diagnoses
During the last few years more than 100 medical secretaries in the Stockholm County have passed a 20 days training course in coding of diagnoses. The medical secretaries are very interested in taking a more active role in the coding. Nevertheless the coding still in most medical departments rests with the doctors who are quite unwilling to transfer the coding to the trained secretaries.

A study using the national patient registers performed by the National Board of Health and Welfare (NBHW) early this year indicates that there is a possible relation between training in coding and the quality of coding.

In order to compare the quality of the coding between medical doctors and specially trained secretaries a study was conducted in 2004 based on 210 medical records from six clinical departments (two from each surgery, orthopaedics and internal medicine). A second and important purpose was to develop a system for continuous in-house quality control performed by the medical secretaries in the clinical departments.

Additional to the general review procedure described earlier coding in this project was done in three steps.

1. Original coding by the clinical doctors (A-coders)
2. Coding by two medical secretaries (B-coders)
3. Coding by expert auditors (C-coders)

When using the C-coders coding as a reference (assumed 100% correctness) it was shown that the A-coders had correct primary diagnosis in 67 % while B-coders were correct in 77 %. Regarding secondary diagnoses the result was similar. A-coders correct in 54 % and B-coders in 63 %.
It was also found that there was a difference in the average number of diagnoses coded per case between the three groups of coders. This study thus underlines and strengthens the results from the study carried out by the NBHW.

Number of diagnoses per case

A-coder	2.6
B-coder	3.6
C-coder	3.0

It is apparent that A-coders have a tendency of down coding (registrating less codes than correct) while the opposite (up coding) relates to the trained secretaries.

During the last PSC/E meeting in Budapest in 2004 a model presented by dr Michael Wilke which offers possibilities to compare quality in coding in a different and maybe more sophisticated model than what has been previously used in Stockholm. The system is in short described below.

Wilkes scoring system

Primary diagnoses	Correct	20 points
	Minor error	15 points
	Incorrect	0 points
Secondary diagnoses	Correctness	0-5 points
	Completeness	Minus 5 -+ 5 points
	Up-coding	Minus 1 point /diagnosis

Wilkes model also includes a scoring module for procedures which was not used in the Stockholm study.
Applying Wilkes scoring system on the data from the quality improvement project results in an average score of 21.7 points for A-coders while B-coders achieve 25.8 points out of a maximum score of 30 points. This indicates that trained secretaries have a much higher quality in coding than the average clinical doctor.

It would be of interest to further develop a consensus about an international scoring system to be used for developing the quality of coding.
It will also be a necessity to have enough skilled coders in order to describe the activities in the health care sector in a correct and plausible way. This is not elaborated in this paper but will be a challenge for the future.